



GÖTEBORGS
UNIVERSITET

Diagnostiska läs- och skrivprov

Bedömning som formar lärandet

Maria Loustalot-Forest
Specialpedagogiska programmet



Namn: Maria Loustalot-Forest
Uppsats/Examensarbete: 15 hp
Kurs: SPP 600
Nivå: Avancerad nivå
Termin/år: VT 2019
Handledare: Ernst Thoutenhoofd
Examinator: Per-Olof Thång

Nyckelord: Diagnostiska läs- och skrivprov, DLS, formativ bedömning, formativ utvärdering, formell lärandet, prov, läs- och skrivscreening, läsprov

Abstract

DLS - diagnostiska läs- och skrivprov är ett frekvent använt screeningmaterial som köps in av skolor för att kartlägga elevers läs- och skrivförmåga på grupp- och individnivå. Det är normrelaterade, standardiserade prov som tillåter jämförelser eftersom proven har prövats ut på en stor grupp jämnåriga. Resultaten ger en god överblick över läs- och skrivfärdigheten i en klass eller på en hel skola. Denna fallstudie av kvantitativ karaktär analyserar ca 350 elevers resultat på fyra olika DLS-prov i årskurs 7 och 8, samt undersöker sambandet mellan resultaten och elevernas betyg i tre ämnen. Syftet är att genom denna kvantitativa undersökning diskutera hur informationen från proven kan användas formativt i undervisningen i högre grad än den görs idag.

Studien finner att skolans elever presterar genomsnittligt enligt standardskalan Stanine som används för att jämföra elever, men resultaten visar samtidigt oroväckande låga värden gällande till exempel ordförståelse och läshastighet. Utgången i proven varierar något mellan klasser och det visar sig också att elever med svenska som andra språk presterar sämre än elever som läser svenska. Däremot uppvisar flickor och pojkar likvärdiga resultat. Anmärkningsvärt är också hur få av de elever som presterar lågt på testerna som får den uppföljning och hjälp i sin läs- och skrivutveckling som vore önskvärt.

Ett positivt samband mellan elevers resultat på DLS-proven och deras betyg framträder, vilket garanterar provens diagnostiska validitet. Ju lägre Stanine-värde på proven desto större sannolikhet att eleven får underkända eller låga betyg. Det finns dock elever som avviker från detta mönster med höga Stanine-värden trots låga betyg, vars resultat bör undersökas närmare. Den generella tendensen är dock att DLS-resultat och betyg i svenska, matematik och samhällsorienterande ämnen samvarierar, vilket ger skolor en fingervisning om vilka elever som riskerar skolsvårigheter på ett generellt plan, inte enbart när det gäller läsning och skrivande. Intressant är att alla fyra prov i relation till betygen i ämnena uppvisade liknande positiva samband, vilket pekar på möjligheten att enbart använda ett eller två digitala prov i diagnostiskt syfte, istället för fyra. En utveckling av analys och uppföljning av DLS-resultaten på individ-, grupp- och skolnivå är angeläget på den undersökta skolan, för att minska risken för skolmisslyckande. Att utgå ifrån begreppet formativ bedömning i analysarbetet av DLS-resultaten är önskvärt eftersom detta kan vägleda lärare i att forma undervisningen för att utveckla elevernas lärande. Specialpedagogerna på skolan bör starta upp detta arbete och vägleda lärare i att utveckla en formativ praktik med DLS som bedömningsinstrument.

Innehållsförteckning

1	Inledning	4
1.1	Bakgrund	7
1.1.1	Begrepp	7
1.1.2	Diagnostiska läs- och skrivprov - DLS	8
1.1.3	Deltesterna	11
1.1.4	Betygen i den svenska grundskolan	13
1.1.5	Särskilt stöd	14
1.1.6	Svenska som andra språk	15
1.1.7	Begreppet formativ bedömning	15
1.2	Problemformulering	17
1.2.1	Syfte	18
1.2.2	Frågeställningar	19
1.2.3	Avgränsningar	19
2	Tidigare forskning	20
2.1	Forskning om läs- och skrivscreening	20
2.2	Forskning om DLS-tester i relation till olika variabler	22
2.3	Forskning om formativ bedömning	24
2.3.1	Formativ begreppsutveckling	24
2.3.2	Skillnaden mellan summativ och formativ bedömning	26
2.3.3	Formativ bedömning på policynivå	28
3	Teorianknytning	30
3.1	Teoretisk ram	30
3.1.1	Psykometrisk tradition	31
3.1.2	Kognitivism	33
3.1.3	Socialkonstruktivistiska/sociokulturella perspektiv	35
3.1.4	Teoretisk utgångspunkt	36
4	Metod	38
4.1	Studiens metod	38
4.2	Urval	38
4.3	Tillvägagångssätt	39
4.3.1	Studiens trovärdighet	40
4.3	Etik	42
5	Resultat	43

5.1	Elevernas resultat	43
5.1.1	Resultatet för elever som läser svenska och svenska som andra språk.....	46
5.1.2	Skillnader klassvis	48
5.1.3	Förhållandet mellan låga resultat på DLS-proven och särskilt stöd	49
5.1.4	Sambandet mellan resultat på DLS-proven och betyg	50
6	Diskussion	54
6.1	Resultatdiskussion	54
	Elevernas resultat och formativ användning av DLS.....	54
	Elever med svenska som andra språk	56
	Elever i behov av särskilt stöd	56
	Sambandet mellan resultat och betyg.....	57
	Hinder och möjligheter för DLS-provens användande.....	59
6.2	Metoddiskussion	61
6.3	Studiens kunskapsbidrag.....	62
6.4	Förslag till vidare forskning	63
7	Referenslista.....	64

Bilagor

Figurförteckning

Figur 1. Stanineskalan	10
Figur 2. Antal elever per Stanine-värde på rättstavningsprovet	44
Figur 3. Antal elever per Stanine-värde för läshastighetprovet	45
Figur 4. Antal elever per Stanine-värde på ordförståelseprovet	45
Figur 5. Antal elever per Stanine-värde på läsförståelseprovet	46
Figur 6. Resultatfördelning elever som läser svenska och svenska som andra språk för provet rättstavning.	46
Figur 7. Resultatfördelning elever som läser svenska och svenska som andraspråk ordförståelseprovet.	47
Figur 8. Andel elever som får särskilt stöd relaterat till resultatet i läshastighetstestet	49
Figur 9. Andel elever som får särskilt stöd relaterat till resultatet i ordförståelseprovet	50
Figur 10. Låddiagram över sambandet mellan elevernas betyg i svenska och resultat på rättstavningsprovet	51
Figur 11. Sambandet mellan elevernas betyg i matematik och resultat på rättstavningsprovet	52
Figur 12. Sambandet mellan elevernas betyg i historia/samhällskunskap och resultat på rättstavningsprovet.	53
Figur 13. Skillnaden i resultat mellan elever som läser svenska och svenska som andra språk, vid provet läshastighet.	74
Figur 14. Skillnaden i resultat mellan elever som läser svenska och svenska som andra språk, vid provet läsförståelse.	74
Figur 15. Andel elever som får särskilt stöd sett till resultatet i rättstavningsprovet	75
Figur 16. Andel elever som får särskilt stöd sett till resultatet i läsförståelseprovet	75
Figur 17. Sambanden för alla fyra proven och betyg i svenska och svenska som andra språk	76
Figur 18. Sambanden för alla fyra proven och betyg i matematik	77
Figur 19. Sambanden för alla fyra proven och betyg i matematik	78

Tabellförteckning

Tabell 1. Antalet elever, antal pojkar och flickor, antal elever som läser svenska eller svenska som andra språk per prov.	39
Tabell 2. Fördelning av elevernas Stanine-värden i procent per prov	43
Tabell 3. Klassernas genomsnittliga Stanine-värden vid respektive prov	48

1 Inledning

Att mäta och bedöma kunskaper och färdigheter är en nödvändig del av vardagen i skolan (Hirsh, 2017). Ett vanligt sätt att mäta elevers kunskaper är prov av olika slag, även om andra typer av bedömningar som mer inriktar sig på kvaliteten av elevers prestationer också har utvecklats (Jönsson, 2018). Oavsett vad vi tycker om prov är det ett faktum att de förekommer, framförallt för att pröva elevers kunskaper gentemot kunskapskraven i olika ämnen (Wikström, 2013). Andra vanliga prov är nationella prov som ska underlätta en jämförelse mellan elever och skolor nationellt samt nationella bedömningsstöd som Skolverket ger ut med syfte att kartlägga elevers kunskaper gentemot målen (Skolverket, 2019c). Var tredje år görs också storskaliga internationella kunskapsprov till exempel OECD:s mätning PISA (*Programme for International Student Assessment*) där 15-åringar i 72 länder testas i läsning, matematik och naturvetenskap. Resultaten har belyst svenska elevers försämrade förmåga inom alla tre testområden (Skolverket, 2016) och anses som ett bättre mått på elevers kunskaper än de nationella proven (SOU 2016:59) då de ger möjlighet att fastställa elevers kunskaper över tid.

Ytterligare en typ av prov som blivit allt vanligare i den svenska skolan är så kallade läs- och skrivscreeningar, med syfte att undersöka stora elevgruppers läs- och skrivförmåga (Myrberg & Lange, 2006; SOU 2016:59; Johansson, 2016). Det är standardiserade, normbaserade tester i till exempel läshastighet, läsförståelse, ordkunskap och stavning för en hel årskurs eller en grupp utformade av privata företag i samarbete med läs- och skrivexperter (Järpsten & Taube, 2018) som köps in av skolor och kommuner. Prov som mäter läs- och skrivutvecklingen hos elever kan sägas ha blivit vanligare sedan decentraliseringen av skolan på 1990-talet då den statliga styrningen minskade (Myrberg & Lange, 2006; Wikström, 2013; Korp, 2011) och intresset för dessa prov har ökat i takt med att ungas läs- och skrivförmåga har kommit att diskuteras allt mer i den offentliga debatten. Den s. k. ”vänta-och-se-pedagogiken” (Johansson, 2016, s.10) har gått tillbaka och numera arbetar man med kartläggning och noggrann uppföljning på många skolor och i många kommuner, menar Johansson. Ett vanligt förekommande prov som skolor använder för läs- och skrivscreening är Diagnostiska läs- och skrivprov (DLS) från Hogrefe Psykologiförlaget AB (Hogrefe Förlag AB, 2019) som ska bli föremål för denna undersökning.

DLS-testerna ingår ofta som en del i screeningplanen hos flera kommuner till exempel Sundsvalls kommun (2016), Täby kommun (2016) och Västerås stad (2016) för att bara nämna några. Friskolekoncernen Academedia har använt DLS som kartläggningsmaterial i sin så kallade "läsa-skriva-räkna" garanti på runt 80 grundskolor (SOU 2016:59). 80 % av en stor del av skolhuvudmännen i Sverige, har i en enkätundersökning redovisat att de genomför egna kartläggningar för att följa upp elevers läs- och skrivutveckling (SOU 2016:59). Detta indikerar att DLS och liknande kartläggningsmaterial används flitigt i svenska skolor. Det är vanligt att specialpedagoger och speciallärare är involverade i administrering, rättning och analys av DLS-proven, men även lärare involveras i testprocessen.

Prov i skolan ger information som kan användas på flera sätt, till exempel summativt där elevens kunskaper för ett kunskapskriterium summeras i efterhand med en siffra eller en bokstav, som sedan får konsekvenser för elevens betyg och framtid. En annan funktion ett prov kan ha är formativt då resultatet används för att ge eleven relevant feedback, att förändra undervisningen eller på annat sätt utveckla lärandet medan det fortgår (Wikström, 2013; Lundahl, 2014; Jönsson, 2018). Summativ bedömning kan definieras "bedömning av lärande", då den summerar elevens kunskaper medan formativ bedömning benämns "bedömning för lärande" då den används för att förbättra och utveckla lärandet (Gustafsson, Cliffordson & Ericksson, 2014; Hirsch, 2017). Medan summativa prov har dominerat skolan och fortfarande är vanliga (Wikström, 2013; Jönsson, 2018) har forskning visat att formativ bedömning främjar lärandet (Black & Wiliam, 1998; Hattie, 2009; Hirsch & Lindberg, 2015; Jönsson, 2018; Lundahl, 2014). Skolverket arbetar sedan länge aktivt med att förmedla kunskap om formativ bedömning och har givit ut flera skrifter med stöd, råd och rekommendationer om hur lärare kan använda formativ bedömning (Skolverket, 2011a; Skolverket 2011b; Hirsch, 2017).

DLS-proven är normerande prov med syfte att differentiera elever kunskapsmässigt och jämföra dem med varandra (Järpsten & Taube, 2018), men provens funktion är helt och hållet formativ. Testerna används inte som betygsunderlag, utan ger istället information om vem som kan ha svårigheter med läsning och skrivning, hur undervisningen ska planeras eller utvärderas i olika ämnen eller hur resurser ska fördelas (Järpsten & Taube, 2018; Roe, 2015). DLS ska fungera som ett pedagogiskt instrument för lärare, specialpedagoger och skolledare för att utveckla elevernas lärande på olika sätt (individuellt eller i grupp), men inte som sorteringsmekanism (Roe, 2015; Fridolfsson, 2015). Rätt använda kan DLS-proven ge både elever och skolpersonal en god och relativt tillförlitlig översikt över elevernas nivå när det

gäller grundläggande färdigheter som läsning, läsförståelse, ordförståelse och stavning (Järpsten & Taube, 2018).

Som specialpedagog har DLS varit det screeningmaterial jag arbetat allra mest med, men min uppfattning är att testresultaten inte används ett formativt sätt som främjar lärande hos eleverna, men inte heller summativt som ett sätt att rangordna eller sortera elever. Lärare kan uttrycka att proven ”inte säger så mycket” om elevernas förmåga, utan ses som ett specialpedagogiskt instrument, skilt från skolans dagliga verksamhet med planering av undervisningen och betygsättning.

Istället fokuserar både lärare, specialpedagoger och skolledning framförallt på elevernas betyg, betygsutveckling och resultat på nationella prov, så kallade summativa bedömningar. Enligt Gustafsson, Cliffordson, & Erickson (2014) och Klapp (2016) är betygen i Sverige av mycket stor betydelse för eleven och dennes framtid, självbild och motivation, så kallade *high-stakes*. De är också viktiga för skolor då statistik årligen redovisar hur många elever som går ut årskurs 9 med godkända betyg (minst E). När skolan är marknadsorienterad och skolors ekonomi är knutna till elevantal blir andel elever med godkända betyg en del i en skolas marknadsföring (Vlachos, 2018). Betygsdominansen i kombination med den begränsade användning av DLS som jag sett på flera skolor väckte mitt intresse för detta ämne och jag vill nu fördjupa mig i hur vi kan använda informationen från dessa omfattande screeningar på ett sätt som främjar elevernas lärande.

DLS-provens användande på en specifik skola och vad proven mäter är således i fokus för den fortsatta framställningen. Jag kommer att presentera resultaten från fyra DLS-prov gjorda under läsåret 2018/2019 på skolan där jag arbetar. Jag presenterar hur dessa skiljer sig åt mellan klasser, mellan elever som läser svenska respektive svenska som andra språk samt hur stödet ser ut för de elever som får lägst poäng på proven. Därefter analyseras huruvida det finns ett samband mellan DLS-resultat och elevernas betyg i svenska/svenska som andra språk, matematik och samhällskunskap/historia. Ett eventuellt samband mellan DLS-provens resultat och betygen, inte bara i svenska, utan i matematik och samhällskunskap, skulle kunna bli ett incitament för lärare och skollära att använda screening på ett sätt som innebär ökat lärande för eleverna. Utifrån teori och forskning diskuterar jag också hinder och möjligheter för att använda DLS-resultaten formativt på den aktuella skolan.

1.1 Bakgrund

Här presenteras termer och begrepp som har betydelse för förståelsen av undersökningen, bland annat kommer DLS-provens innehåll och begreppet formativ bedömning att förklaras. Därefter följer problemformulering och syfte.

1.1.1 Begrepp

Screening

DLS bör helst genomföras i grupp för screeningsyfte, menar Järpsten & Taube (2018). Enligt Nationalencyklopedin är screening en metod att sälla fram till exempel individer med sökt egenskap i en stor grupp (Nationalencyklopedin, 2019). Med "läs- och skrivscreening" menar vi alltså en provning på gruppnivå med syfte att sälla fram individer som ligger under den förväntade nivån när det gäller läsning och skrivning. Informationen som insamlas under en sådan screening kan användas för formativ bedömning (Korp, 2011, Jönsson, 2018).

Mätning och bedömning

Två andra begrepp som används i texten är mätning och bedömning. Mätning används när vi vill fastställa storleken eller omfattningen på något och kvantifiera detta. När vi gör en bedömning så värderar vi vad mätningen visade i relation till något annat (Wikström, 2013). Jönsson (2018) menar att bedömning kan betyda många saker. Det kan användas som ett övergripande begrepp där prov ingår som en typ av bedömning, men också syfta på processen eller utfallet som i ett omdöme eller ett provresultat. Jönsson håller med Wikström att bedömning inte är en mätning och därmed inte är kvantitativ till sin karaktär, utan handlar mer om att bestämma kvaliteten på en uppgift (Jönsson, 2018).

Test och prov

Begreppen *test* och *prov* förekommer med samma betydelse i denna text, även om prov är mer vedertaget i svenskan med avseende på kunskapsmätningar i skolan (Wikström, 2013). Jönsson (2018) definierar prov som "ett urval av uppgifter, som tillsammans används för att mäta individers kunskaper" (s. 11). Prov är ett sätt att "synliggöra människors kunskaper, vilket innebär att man låter dem svara på ett antal uppgifter" (Jönsson, 2018, s. 13) och ju fler rätta svar eleven har desto bättre kunskaper anses eleven ha.

Ett prov som undersöker kunskaper i relation till ett visst kunskapsstoff benämns målrelaterat eller kriterierelaterat, men prov som ges för att klassificera eller jämföra provtagarna kallas normrelaterade. Högskoleprovet är ett exempel på ett normrelaterat prov. Resultatet tolkas utifrån hur samtliga provtagare presterat och redovisas på en standardiserad och normerad skala från 0,0–2,0. Ett normrelaterat prov handlar helt enkelt om att bedöma resultatet inom gruppen (Wikström, 2017). Wikström (2017) menar att summativa och formativa prov inte är två typer prov utan två olika tolkningar av provresultatet, som kompletterar varandra.

Reliabilitet och validitet

I test och provsammanhang används begreppen reliabilitet och validitet när ett provs mätsäkerhet och giltighet diskuteras. Reliabiliteten handlar om hur tillförlitligt ett mätresultat är eller hur stabil mätningen är. Ett prov med hög reliabilitet ger samma resultat om provet genomförs många gånger om (Wikström, 2013; Gustafsson m.fl., 2014). Tillförlitlighet avgörs också av hur provet är sammansatt till exempel om det är för få frågor, men också om elever är sjuka, har bråttom, är helt oförberedda, väldigt nervösa, inte känner igen provformatet eller inte hinner göra klart provet (Wikström, 2013). Validitet definieras av Jönsson (2018) som ” att svara på frågan om ett prov – eller en bedömning – testar det som provet/bedömningen var avsedd att testa” (s. 55) eller huruvida provet mäter ”rätt sak på rätt sätt” enligt Wikström (2013, s. 44). Ett provs validitet påverkas bland annat av vilket syfte provet har, vilket samband resultatet har med det provet vill mäta och hur resultaten analyseras (Jönsson, 2018; Wikström, 2017). Andra sätt att undersöka validiteten är att undersöka hur väl provresultatet korrelerar med ett uppsatt kriterium till exempel betyg eller nationella prov, vilket sker i denna undersökning (Elbro & Poulsen, 2016), vilket också delvis tangerar uppsatsens ämne. DLS-provens validitet och reliabilitet diskuteras mer utförligt i kapitel 4. Nedan följer en redogörelse för DLS-proven.

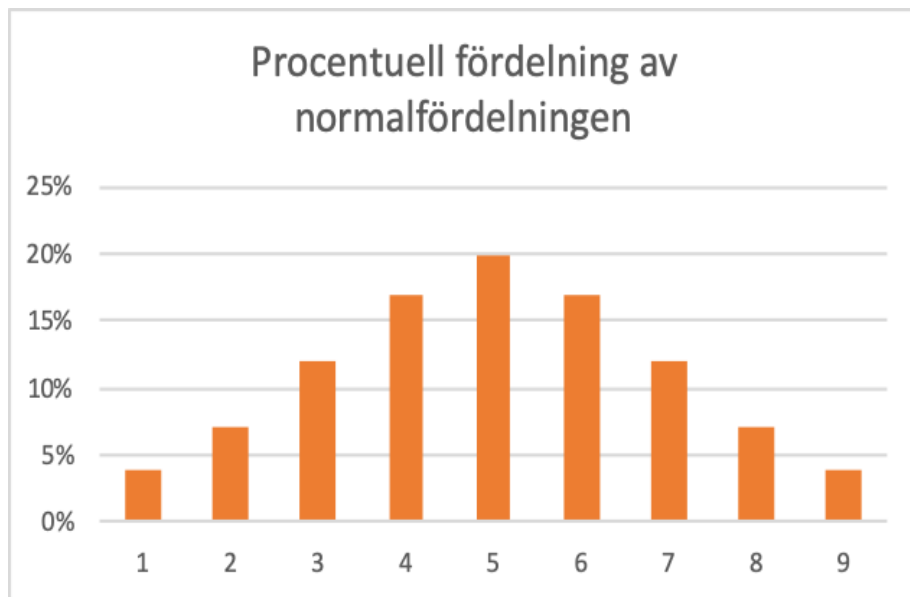
1.1.2 Diagnostiska läs- och skrivprov - DLS

Diagnostiska läs- och skrivprov (DLS) är ett vedertaget material för att kartlägga grupper och enskilda elevers färdigheter i läsning och skrivning från 8–16 års ålder (Järpsten & Taube, 2018; SPSM, 2019; SBU, 2014). Wikström (2017) menar att det tar lång tid att utveckla den här typen av storskaliga prov, många månader av arbete med granskningar, utprovningar och kravgränssättningsarbete.

År 2018 reviderade Psykologiförlaget Hogrefe förlag DLS-proven rättstavning, ordförståelse, läshastighet och läsförståelse för skolåren 7–9 och år 1 på gymnasiet som tidigare konstruerats och standardiserats under 2001–2002 (Järpsten, 2002). Av dessa reviderade prov från 2018 är proven läshastighet och ordförståelse i digital form och proven i stavning och läsförståelse i pappersform.

Enligt handledningen för DLS har utgångspunkten för provkonstruktörerna varit vad läroplanen för grundskolans Lgr11 (Skolverket, 2018) säger om ämnet svenska och författarna har vid provkonstruktionen utgått från syftet med ämnet, målsättningar, centralt innehåll samt uppnåendemål i slutet av årskurs 9 för att uppnå innehållsvaliditet, det vill säga att provet ska kunna vara exempel på det kunskapsområde det avser mäta (Järpsten & Taube, 2018; Elbro & Poulsen, 2016).

DLS är ett så kallat normrelaterat prov som grundar sig på att det finns olika prestationer som man vill ta reda på hur de fördelar sig. En sådan mätning kallas relativ och innebär att resultatet tolkas utifrån hur gruppen har presterat (Wikström, 2013). Elevens råpoäng på DLS-proven överförs till en standardskala, för att det ska vara möjligt att göra normjämförelser. Resultaten kan tolkas i relation till hur en stor grupp jämnåriga presterade på proven under en utprovningsperiod mellan 2016–2018, sammanlagt 5600 elever. Standardskalan stanine används i DLS-proven; den bygger på en indelning av normalfördelningskurvan i nio skalsteg. Medelvärde är det femte steget på skalan med en standardavvikelse på två skalsteg (4–6). Figuren visar hur normalfördelningen procentuellt fördelar sig över nio skalsteg, värden som ligger nära medelvärdet 5 är vanligare än värden som ligger långt ifrån, därför blir formen på normalfördelningen som en kulle eller en klocka så som figur 1 visar.



Figur 1. Stanineskalen

Från skalsteg 1 till och med 3 faller de 23 % svagaste provresultaten vid standardiseringen och mellan skalsteg 7–9 de 23 % starkaste resultaten. Elever med Stanine-värde mellan 1–3 riskerar svårigheter och behöver följas upp med intervjuer och tester (Järpsten & Taube, 2018). För att jämförelse med normerna ska vara möjlig bör proven rättstavning, ordförståelse och läshastighet genomföras under oktober till november enligt Järpsten och Taube. I denna studie genomfördes proven en månad tidigare, i månadsskiftet augusti/september, vilket påverkar tillförlitligheten något. Eleverna hade då precis börjat i skolan efter sommaruppehållet. Proven i läsförståelse genomfördes i februari och mars enligt rekommendationen.

1.1.3 Deltesterna

Rättstavning

Syftet med rättstavningsprovet är att få en uppfattning om elevernas förmåga att stava vanliga ord, vilket finns med under rubriken ”Centralt innehåll i årskurs 7–9”, Läsa och skriva” (Skolverket, 2018) i följande formulering: ”Språkets struktur med stavningsregler, skiljetecken, ordklasser och satsdelar” (s. 261).

Det finns enligt Fridolfsson (2015) ett starkt samband mellan avkodning av text och stavning och det är därför viktigt att ha med stavningstest tillsammans med prov i avkodning och läsförståelse när man prövar elevers läs- och skrivförmåga. Fridolfsson rekommenderar att fortsätta pröva elevers stavningsfärdighet också upp i åldrarna.

Testet rättstavning prövar 36 ord som dikteras av läraren insatta i meningar. Majoriteten av orden prövar dubbelteckning och ljudstridighet, men också ljudenligt stavade ord ingår. När testet är genomfört räknas antalet korrekt stavade ord tillsammans i en råpoäng som ligger till grund för normjämförelsen; max 36 poäng. På baksidan av testet finns ett analyschema för att systematisera elevernas stavfel (Järpsten & Taube, 2018). Detta analyschema berörs inte i denna framställning.

Ordförståelse

Med detta test bör läraren få en uppfattning om elevernas generella förmåga att förstå ord och uttryck hämtade ur texter elever möter i skolan. Detta lyfts också fram i läroplanen (Skolverket, 2018) i ämnets centrala innehåll. Under rubriken ”Centralt innehåll i årskurs 7–9, språkbruk står bland annat att eleven ska känna till ord och begrepp, nya ord och låneord i språket samt dess nyanser och värdeladdning (Skolverket, 2018, s. 261–262).

34 ord ingår i varsin fras där eleven ska markera ett svarsalternativ av fyra möjliga efter varje ord. Orden tillhör olika ordklasser och hämtas ur skönlitteratur såväl som sakprosa. Råpoängen räknas ihop utifrån antalet korrekt löst uppgifter, maximalt kan en elev få 34 poäng (Järpsten & Taube, 2018). Taube, Fredriksson och Olofsson (2015) menar i sin forskningsöversikt att bristande ordförståelse ofta hindrar läsförståelse.

Läshastighet

Här prövas elevernas förmåga att läsa snabbt med bibehållen förståelse, elevernas läsflyt. En sammanhängande text innehållande 36 parenteser eller distraktorer läses tyst. Dessa distraktorer innehåller tre alternativ varav ett passar in i sammanhanget och framgår tydligt. Detta ska visa hur långt eleverna har hunnit läsa med förståelse under lästiden. Eleven väljer det ord i parentesen som passar bäst in i sammanhanget. Det antal ord eleven har strukit under korrekt blir råpoäng som ligger till grund för normjämförelsen. Här kan också sambandet mellan elevens läshastighet och elevens läsförståelse observeras.

Läsförståelse

Det övergripande målet med all läsning är att förstå det man läser, men det är inte förrän avkodningsförmågan är helt etablerad som det är meningsfullt att pröva läsförståelsen (Fridolfsson, 2015). Samtidigt kan det vara komplicerat att pröva läsförståelse hos elever eftersom elevens motivation, intresse och erfarenheter av olika ämnen som förekommer i testerna också påverkar resultatet, menar Fridolfsson, och rekommenderar återkommande tester för att pröva läsförståelse hos elever, något som också förekommer på den aktuella skolan.

I DLS-provet av elevernas läsförståelse lyfts elevernas användning av effektiva lässtrategier och förståelsen för olika typer av texter fram. I Lgr11 under rubriken ”Centralt innehåll i årskurs 7–9, under rubriken ”Läsa och skriva” fastställs att elever ska undervisas i lässtrategier för att förstå texters budskap samt att lära sig urskilja texters budskap (s. 261).

Provet i läsförståelse ska ge information om elevernas förmåga att läsa och förstå olika typer av texter samt själv formulera iakttagelser, tolkningar och slutsatser utifrån texten. Utgångspunkten för läsförståelseprovet har varit PISA-undersökningens beskrivning av olika läsprocesser. Frågorna i provet mäter lässtrategierna informationssökning, tolkning och reflektion, men för att lösa en uppgift behövs egentligen alla tre strategierna.

Testet har tre sammanhängande texter med tillhörande uppgifter. Version A innehåller totalt 24 uppgifter, varav 21 är flervalsuppgifter och eleven markerar en av dem. Tre uppgifter kräver öppna svar och mäter elevernas förmåga reflektera över den lästa texten.

Det är antalet korrekt besvarade flervalsuppgifter som ligger till grund för normjämförelsen. Det är viktigt att också göra en kvalitativ utvärdering av elevernas läsförståelse, särskilt för de

elever som hamnat i riskzonen för läs- och skrivproblematik, med Stanine-värden mellan 1-3. (Järpsten & Taube, 2018).

DLS provens teoretiska grund

Enligt Järpsten & Taube (2018) vilar DLS proven på en allmän accepterad formel för läsförståelse *the simple view of reading*, definierad av Gough & Turner (1986): Läsförståelse = avkodning x förståelse av språk. Den innebär att förmågan att förstå skriven text beror på om individen kan avkoda, se vad som står, snabbt och rätt samt kan förstå talat språk. Om eleven inte har en snabb och effektiv avkodning påverkas förståelsen negativt. När en god muntlig förståelse saknas påverkar detta också läsförståelsen i skriven text. Hoover & Tunmer (2018) visar i sin översikt av tre forskningsstudier att *the simple view of reading* försätter att vara en robust modell och relevant när det gäller att förklara svårigheter i läsförståelse.

DLS grundar sig likafullt på en liknande teori kring begreppet skrivande: Skrivning = stavning x ideation/tankebyggande (Juel, Griffith & Gough, 1986). I artikeln från 1986 visar författarna att fonetisk medvetenhet är grunden för att utveckla stavning, men att skrivandet också är beroende av att eleven kan generera idéer till sitt skrivande, snarare än god läsförståelse.

Enligt Järpsten & Taube (2018) är läsning och skrivande tätt sammanlänkade. I början av grundskolan domineras undervisningen av den mer tekniska sidan av läsning och skrivande och därefter tar den innehållsmässiga sidan av läsning och skrivande över. I DLS-proven rättstavning och läshastighet mäts den tekniska sidan av stavning och läsning, medan ordförståelseprovet och läsförståelseprovet ger oss mer information om den innehållsmässiga sida. Det faktum att två av proven görs digitalt anses inte påverka mätsäkerheten eftersom den digitala versionen överensstämmer med pappersversionen (Järpsten & Taube, 2018).

1.1.4 Betygen i den svenska grundskolan

Ett betyg är en skriftlig handling som rymmer ett omdöme gällande till exempel en persons kunskaper, förmågor och kvalifikationer. Betyg förutsätter en jämförelse, i nuläget med kunskapskrav definierade i läroplanen (Gustavsson, Måhl & Sundblad, 2012). Det rör sig om en summativ bedömning för att pröva om en elev har uppnått de kunskaper som eftersträvas under en relativt lång studieperiod som till exempel en kurs, en termin, ett läsår eller 9-årig

grundskola (Korp, 2011). Detta skiljer sig från de pågående bedömningar (formativ bedömning) som lärare kan använda sig av för att ge feedback och förändra undervisningen som kan stärka elevers lärande (Gustafsson m.fl., 2014).

Betyg ges i slutet av varje termin från och med årskurs 6 (SFS 2010:800). Alla kunskapskrav för en fastställd betygsnivå ska vara uppfyllda för att kunna få ett visst betyg. Det finns kunskapskriterier för betygsnivåerna A, C och E. Alla kunskapskrav som beskriver en nivå behöver vara uppfyllda för att eleven ska få betyget E, det räcker inte att nå upp till några kriterier. För betygen D och B finns inga rubricerade kunskapssteg, för dessa behöver den undre nivån till exempel E vara helt uppfyllda och kraven för den övre nivån till exempel C vara uppfyllda till ”övervägande del”.

Betygens roll

Jönsson (2018) menar att betyg har tre ändamål: information, motivation och urval. Det är förvisso en viktig information till elever och föräldrar, men en alltför begränsad för att den ska leda till en förändring i lärandet hävdar Jönsson. Betygen fungerar inte heller som motivator för elever med låga betyg, snarare visar det sig att elever då ser betyget som ett kvitto på att de är ”dumma” och slutar studera. Som urval har betygen en viss betydelse, särskilt i slutet av grundskolans årskurs 9, men svårigheten med att tolka betygskriterierna gör att likvärdigheten hotas (Jönsson, 2018). Gustafsson m.fl. (2014) kommer i rapporten om likvärdig kunskapsbedömning fram till att betygssystemet brister i likvärdighet och rättssäkerhet. Utredningen Samling för skolan (SOU 2017:35) pekar på vikten av godkända betyg från grundskolan som en skyddsmekanism för ungdomar när det gäller att inte hamna i kriminalitet, missbruk och bidragsberoende. Kunskapskraven för betyget E (Godkänt) kan anses som grundläggande nivå för att individen ska fungera som samhällsmedlem. Det ideala förhållandet är att alla elever med skolplikt ska nå godkända betyg i alla ämnen när de lämnar årskurs 9, menar utredningens författare (SOU 2017:35).

1.1.5 Särskilt stöd

Elevers rätt till stöd regleras i skollagen (SFS 2011:800), men det är inte reglerat hur lärare ska upptäcka elevernas behov, vilka tester de ska använda eller hur de ska tolka betyg, utan läraren bedöms kunna slå larm utifrån sin yrkesmässighet. Däremot är det tydligt att rektor

har det yttersta ansvaret för att eleverna får det stöd de har rätt till. När en elev riskerar att inte nå målen ansvarar rektor för att en utredning genomförs. Utredningen ska visa vilka åtgärder som behöver sättas in och om ett åtgärdsprogram behöver upprättas (Lundhal, Hultén & Tveit, 2016).

1.1.6 Svenska som andra språk

Svenska och svenska som andra språk (sva) är likvärdiga ämnen inom skolan och ger samma behörighet till gymnasiet och högre utbildning. För att ha rätt att läsa sva ska eleven antingen ha ett annat modersmål än svenska, ha gått i skola i utlandet eller tala svenska med en förälder. Kursplanerna är relativt lika, men lärarutbildningarna skiljer sig åt avsevärt mellan svenska och svenska som andra språk. Elever som har behov att läsa svenska som andra språk ska få göra detta oberoende vilken språklig nivå de befinner sig på. Eleverna i sva har rätt att få en undervisning som tar hänsyn till att det är ett andraspråk som utvecklas, enligt Nationellt centrum för svenska som andra språk (2019).

1.1.7 Begreppet formativ bedömning

Formativ bedömning är ett begrepp inom bedömningsforskning som fått ökat inflytande tack vare en forskningsöversikt av Paul Black och Dylan Wiliam (1998). Den har vuxit fram som en reaktion på den summativa bedömning som har dominerat och fortfarande dominerar skolväsendet. Med en summativ bedömning av en elevs kunskaper menas bedömning som prövar vad eleven har lärt sig eller om ett visst mål är uppnått och som sedan kan ligga till grund för olika slags beslut kopplade till individen och organisationen, till exempel ett betyg eller inträde till högre utbildning (Lundahl, 2014). De prov som används summativt är de vanligast förekommande proven i skolan och i samhället, som de flesta är vana vid (Wikström, 2013). Den formativa bedömningen, å andra sidan, är till exempel prov eller andra uppgifter med ändamål att ge läraren och eleven information om elevens befintliga kunskapsnivå som vägledning för att göra förändringar i undervisningen eller som stöd för eleven som kan använda informationen för att förbättra sitt lärande. När prestationen eller utfallet används för att anpassa eller ”forma” något så används det alltså på ett formativt sätt, till exempel ett diagnostiskt prov innan en kurs startar som vägleder läraren om hur denne ska planera undervisningen lika gärna som att resultatet från det diagnostiska testet eller återkopplingen från bedömningen kan vägleda den individuella studenten i vilket arbete som

krävs för att klara målen för kunskapsområdet (Hirsh 2017; Dixson & Worrell, 2016; Lundahl, 2014; Hassan, 2011). Viktigt att understryka är att ett prov som sådant inte är formativt eller summativt, utan att resultaten kan användas i summativt eller formativt syfte eller både och (Jönsson, 2018; Wikström, 2013; Skolverket 2011b, Korp, 2011).

Bakgrund till begreppet formativ bedömning

Trots att idéerna om formativ bedömning har funnits sedan 1930-talet är det Scriven (1967), Bloom, Hastings, & Madaus (1971) & Sadler (1989) som anses ha lagt grunden för den syn på formativ bedömning vi har idag. Scriven (1967) tar upp begreppet formativ *evaluation* (utvärdering) i relation till högre utbildning och hur implementeringen av denna kan användas för att utveckla kursplanen (curriculum). Bloom m.fl. (1971) utvecklar Scrivens (1967) tankegångar och ser formativ *evaluation* (utvärdering) som något som inte bara kan utveckla kursplanen utan som också kan användas i undervisning och elevers lärande.

Sadler (1989) använder istället för *evaluation* (utvärdering) begreppet *assessment* (bedömning) och utvecklar också idén om feedback. Enligt Sadler är feedback en grundläggande del i formativ bedömning, ett kvitto på hur framgångsrikt en elevs prestation är, som ger information både till läraren och eleven. Pedagogen ska använda feedback för att förändra utbildningens upplägg på ett sätt som gagnar elevens lärande ännu mer och eleven ska ta till sig feedback för att uppmärksamma sina egna styrkor och arbeta med sina svagheter menar Sadler (1989).

I Black & Wiliams (1998) tongivande översikt: *Assessment and classroom learning* gör forskarna en bred definition av begreppet formativ bedömning i linje med Sadlers; ”all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (s. 7-8). Enligt denna definition är bedömning i formativt syfte när elevers styrkor och utvecklingsbehov synliggörs och ger lärare eller elever möjlighet att modifiera undervisning eller lärandeaktiviteter (Skolforskningsinstitutet, 2018). Det är denna definition av Black & Wiliams om formativ bedömning som har kommit att ligga till grund för de flesta verksamma forskares, lärares och skolledares syn på formativ bedömning i Sverige och internationellt. Skolverket kom i samband med implementeringen av senaste läroplanen Lgr11 ut med stödmaterial för skolor (till exempel Kunskapsbedömning i skolan, 2011;

Kunskapsbedömning – var, hur och varför, 2011) där Black & Williams syn på formativ bedömning beskrivs och rekommenderas för lärare i grundskolan och gymnasiet.

Begreppet formativ bedömning har ifrågasatts av flera forskare, som varande alltför ospecifikt. Bland annat Broadfoot m.fl. (1999) skriver att termen formativ i sig själv är öppen för många tolkningar och betyder sällan mer än att bedömning sker ofta och är planerad på samma gång som undervisningen planeras. De föreslog istället termen *bedömning för lärande* (Assessment for learning) som började användas i början av 1990-talet. Begreppen används idag som synonymer även om bedömning för lärare ses som något mer övergripande än formativ bedömning. Definitionen som Broadfoot m.fl. (1999) ger är att bedömning för lärare är processen att samla ihop och tolka information som kan användas av lärare och elever för att få reda på var eleverna är i sin lärandeprocess, vart de ska och hur de ska komma dit.

Jönsson (2018) definierar formativ bedömning: ”information från bedömningar används för att stödja elevens lärande” (s. 73) och synliggör den bedömning eller formandet av lärandet som kan ske på gruppnivå när läraren utifrån resultat på prov/bedömningar tar beslut om sin egen undervisning. Korp (2011) anser att denna typ av formativ bedömning på gruppnivå, där individer inte står i fokus, bör definieras som utvärdering istället för bedömning medan termen formativ bedömning framförallt handlar om att ta in och återge kvalitén i elevers prestationer på individnivå (Jönsson, 2018; Korp, 2011).

1.2 Problemformulering

Läs- och skrivscreeningar genomförs på många skolor med hjälp av standardiserade och normrelaterade tester till exempel DLS och ger snabbt och effektivt mycket information om elevers läs- och skrivförmåga (Wolff, 2013). En fördel med digitala hjälpmedel i skolan är att många elever kan bedömas samtidigt (Vetenskapsrådet, 2019). Samtidigt saknas i vissa fall en medvetenhet på skolor omkring hur resultaten ska användas och kopplingarna mellan testerna och de pedagogiska verktygen och insatserna är svaga (Myrberg & Lange, 2006). Arbete med screeningen leds många gånger av elevhälsan och då särskilt den specialpedagogiska professionen, i linje med Socialstyrelsens rekommendationer om att elevhälsan bör arbeta förebyggande och medvetet med bedömning av kunskapsutvecklingen på skolan (Socialstyrelsen/Skolverket, 2016). Resultaten förs ut till lärare och skolledare, men informationen används inte för att styra och omforma undervisningen på det sätt som skulle

kunna vara möjligt. Frågan om elevers läs- och skrivutveckling är angelägen eftersom detta får stor betydelse för att klara skolan och samhällets krav (Reichenberg, 2008).

Pettersson (2010) pekar på att vid all sorts bedömning krävs en medvetenhet bland annat omkring syftet med bedömningen, vilka svagheter och styrkor bedömningen har, om bedömningen ges i formativt eller summativt syfte, vilket enligt min mening sällan är fallet vid användandet av DLS-proven. Myrberg & Lange (2006) menar att testa utan att använda eller återkoppla resultatet för eller till eleverna är slöseri med tid och kan till och med vara skadligt för eleverna.

Att lyfta upp frågan om DLS och resonera kring hur informationen från läs- och skrivscreeningen kan användas och tas tillvara på skolor är angelägen, inte minst när flera bedömningsforskare och Skolverket alla menar att all bedömning ska gagna elevernas lärande (Skolverket, 2011a, Skolverket 2011b; Lundahl, 2014; Jönsson, 2017; Jönsson, 2018; Skolverket 2019a; Skolverket 2019b).

1.2.1 Syfte

Uppsatsens syfte är att på den aktuella skolan undersöka resultat från fyra DLS-prov och vilken relation som finns mellan resultaten och elevernas betyg i svenska, matematik och samhällskunskap/historia höstterminen 2018. Genom att presentera resultaten utifrån olika variabler till exempel de olika testerna, olika klasser, elever som läser svenska som andra språk och elever som är i behov av särskilt stöd vill jag diskutera hur denna information från DLS-proven kan användas formativt på individ-, grupp- och skolnivå samt vilka hinder och möjligheter det finns med detta arbetssätt. Mitt syfte utmynnar i följande frågeställningar.

1.2.2 Frågeställningar

- Hur presterar eleverna i år 7 och 8 i relation till den normerande Stanineskalan i diagnostiska läs- och skrivprov (ordkunskap, läshastighet, stavning och läsförståelse)?
- Hur många av de elever som får Stanine-värde 1–3 har tillgång till särskilt stöd?
- Hur ser relationen ut mellan elevernas betyg i tre olika ämnen och resultaten i de fyra DLS-testerna?

1.2.3 Avgränsningar

Temat för denna studie tangerar flera stora forskningsområden som till exempel elevers läs- och skrivförmåga, bedömning, betygs eller provs reliabilitet och validitet, ungdomars läsförståelse samt insatser, stöd och anpassningar för elever med läs- och skrivsvårigheter. Av utrymmesskäl begränsar jag mig här framförallt till att granska läs- och skrivscreeningen DLS ur ett formativt bedömningsperspektiv, i en teoretisk och historisk kontext. Jag inser vikten av att därefter ta fram pedagogiska insatser på individ-, grupp- och skolnivå (Myrberg & Lange, 2006) med bas i forskning för att förbättra elevers läs- och skrivnivå (Skolverket, 2018b), men detta inkluderas inte i denna text.

Nästa kapitel innehåller en forskningsöversikt i vilken litteratur om läs- och skrivscreening, DLS-proven och formativ bedömning ur olika perspektiv inkluderas.

2 Tidigare forskning

Sökstrategin utformades med utgångspunkt från frågeställningarna och att de skulle vara så breda och objektiva som möjligt. Jag använde och kombinerade sökorden: ”DLS, läs- och skrivprov, läs- och skrivscreening, läs- och skrivkartläggning, kartläggningsmetoder, läshastighetsprov, läsförståelseprov, ordkunskap” och kombinerade dessa med sökorden; ”betyg och bedömning”, ”prov” samt ”formativ bedömning, formativ utvärdering”. Litteratursökningar gjordes framförallt på Göteborgs Universitetsbibliotek (Supersök) från vilken internationella databaser kan nås, men också på Google, Skolverket, Vetenskapsrådet och Uppsatser.se. Utöver detta användes ”snöbollsmetoden”, det vill säga att artiklar och böcker söktes i referenserna från litteratur jag funnit relevant.

Litteraturgenomgången är uppdelad i tre sektioner. Den första berör litteratur om läs- och skrivscreening, den andra om läs- och skrivscreening kopplad till DLS och den tredje tar upp formativ bedömning ur olika perspektiv.

2.1 Forskning om läs- och skrivscreening

Det råder en konsensus bland läsforskare att läs- och skrivscreeningar är viktiga pedagogiska instrument vid professionellt arbete med barn- och ungdomars läsande (Hagtvet, Frost & Refsahl, 2016; Roe, 2015; Fridolfsson, 2015; Lundahl, 2010; Myrdal, 2003). Rätt använda och med respekt för individerna som testas bidrar de till kunskap om elevers nivå med värde för individ, föräldrar och lärandemiljö, menar författarna. Johansson (2016) anser att diagnosticering har tappat mycket av sin ”stigmatiseringsstämpel” när man i högre grad lyfter fram det formativa syftet med att skapa åtgärder, stödinsatser, resurstilldelning samt att anpassa undervisningen. Hagtvet m.fl. (2016) framhäver läs- och skrivtesters formativa värde och menar att det kan ge ”kursjusterande återkoppling till lärare och elev om hur inläringen fortlöper” (s. 178), vilket även övriga läsforskare betonar. (Järpsten & Taube, 2018; Roe, 2015; Fridolfsson, 2015; Johansson, 2016). Johansson (2016) betonar särskilt att eleven och föräldrar bör informeras om hur resultaten ska användas och varför proven görs. Samtalet efteråt med elever som fått låga poäng är viktigt. När det görs på rätt sätt kan skapa positiva möjligheter för eleven att utveckla sin läsning. Myrberg & Lange (2006) för fram att testning och screening ofta förekommer i svenska skolor, men är inte i tillräckligt hög utsträckning

kopplade till pedagogiska åtgärder. En intervjuad skolpsykolog i den så kallade *Konsensusrapporten* uttrycker att ”tester görs, och sedan hänger det i luften på något sätt” (Myrberg & Lange, 2006). Det framkommer i den första och andra *Konsensusrapporten* (Myrberg, 2003; Myrberg & Lange, 2006) där 24 forskare inom läs- och skrivområdet ger sina synpunkter, att normerande läs- och skrivtest bör användas i högre utsträckning än idag och att testpraxis bör organiseras så att skolor kan följa elevers läs- och skrivutveckling över tid. De normerande testerna fyller också en funktion av fortbildning för nytexaminerade lärare som har yttligare kunskap om elevers läs- och skrivutveckling än tidigare generationers lärare, menar forskare inom läsområdet i rapporten (Myrberg & Lange, 2006).

Ett exempel från USA där screeningen fått en central roll och där resultaten följs upp och leder till ökat lärande är i ett vanligt förekommande stödsystem på skolnivå i USA; *Response to Intervention* (Gersten m.fl., 2008). Metoden uppmuntrades i samband med uppdateringen av *Individuals with Disabilities Education Act IDEA 2004* och uppmuntrade delstater att använda RTI för att förebygga lässvårigheter och hitta elever med inlärningssvårigheter. RTI är ett system för att tidigt upptäcka och förebygga till exempel svårigheter i matematik och svenska. Elever på skolor delas upp i tre nivåer. På nivå 1 finner vi eleverna som klarar av att nå läroplanens mål med effektiv och evidensbaserad undervisning, på nivå 2 finns elever som behöver interventioner på gruppnivå för att nå målen och på nivå 3 har vi de elever som trots stöd inte når de uppsatta målen och som behöver individualiserad hjälp. Metoden bygger på regelbunden screening tre gånger per år av alla. RTI är ett exempel på hur screening används formativt på skolnivå. Datainsamlingen skapar ett effektivt och fruktbart samarbete mellan lärare och specialpedagoger (Gersten m. fl., 2008).

Ridgeway, Price, Simpson, & Rose (2011) visar i sin artikel *Reviewing the roots of response to intervention: is there enough research to support the promise?* att det finns evidens bakom de olika delarna av RTI till exempel omfattande skolscreening. Den data som skolorna samlar in i screeningarna kan användas dels för att identifiera elever som ligger efter, men också för att belysa effektiviteten i undervisningen för nivå 1 (alla elever). När 20 % av eleverna inte har förbättrats sina resultat i screeningen ger det signaler att förändra och utveckla innehållet i undervisningen. RTI är en metod där man försökt förena formativ bedömning med screeningförfarandet.

I Utredningen om en läsa-skriva-räkna garanti (SOU 2016:59) går författarna igenom hur ett antal kommuner har organiserat det strukturerade arbetet med kartläggning av elevers läs- och

skrivutveckling samt åtgärder och stöd. Många kommuner har planer, ibland i form av garantier, för kartläggning och för hur man ska ge elever som kommer efter i sin läs- skriv- och matematikutveckling stöd framförallt upp till årskurs 3, i vilka DLS ofta ingår som en del. Åtgärderna är inte lika strukturerade som inom *Response to Intervention*, men utredningen visar att i några kommuner till exempel Kävlinge har det systematiska arbetet givit resultat med fler elever som klarar nationella provet i år 3.

2.2 Forskning om DLS-tester i relation till olika variabler

Ett fåtal svenska artiklar och böcker behandlar specifikt testet Diagnostiska läs- och skrivprov (DLS), dock inte alltid i relation till betyg eller tillsammans med teorin om formativ bedömning. Statens beredning för medicinsk utvärdering (SBU, 2014) nämner DLS som ett test med hög validitet och reliabilitet i sin litteraturoversikt "Dyslexi hos barn och unga".

Flera kvalitativa intervjuundersökningar tar upp specialpedagogers eller lärares syn på och attityder till läs- och skrivscreeningar (Bäckman, 2011; Bylund, 2011; Björnsson & Hellström, 2017). Alla tre kommer i sina texter fram till att läs- och skrivscreeningarna inte förstås fullt ut av lärarna som ser testningen som ett sätt för specialpedagogerna att urskilja vilka som behöver stöd snarare än att de kan ligga till grund för planering av undervisning eller resursfördelning. I Björnsson & Hellströms (2017) uppsats anser de intervjuade speciallärarna på högstadiet att den främsta anledningen till att DLS-testerna används på skolor är att de ingår i en kommunal screeningplan. Testet upplevs som tillförlitligt när det gäller att hitta elever med lässvårigheter, men samtidigt opålitligt eftersom instruktionerna för hur testet ska genomföras inte alltid följs. Bylund (2018) intervjuar 10 speciallärare som anser att det saknas tid för analys av resultaten från DLS samt att insatser för eleverna med lägst poäng inte sätts in. Sammanfattningsvis indikerar de kvalitativa studierna av läs- och skrivscreeningar att det finns utvecklingsmöjligheter när det gäller att använda sig av resultaten på ett formativt sätt, dock är det svårt att dra några generella slutsatser av några få informanter.

Wiktorell (2016) har i en enkätundersökning redogjort för hur en kommunal, årlig läs- och skrivscreening uppfattas av lärare, skolledare och specialpedagoger/speciallärare på skolor i kommunen. Hon kunde se ett generellt motstånd mot läs- och skrivscreeningen bland lärare samt att de lärare som undervisar andra ämnen än svenska inte känner sig berörda av

elevernas resultat på de olika testerna. Screeningresultaten användes främst på individ- och grupp-nivå och inte alls på skol- och kommunnivå. 70 % av skolpersonalen som besvarade enkäten ansåg att man följde riktlinjerna för hur testet skulle genomföras och 60 % av informanterna i studien ansåg att låga prestationer i screeningen följdes upp på något sätt. Wiktorrell drog ändå slutsatsen att resultaten inte användes tillräckligt av lärare, skolledare och kommunen.

Carlsson (2017) undersöker i sin studie orsakerna till att elever i år 7 hade en lägre nivå än förväntat på DLS läsförståelsetest i augusti 2016 samtidigt som alla elever klarat målen för läsförståelse i de nationella proven i svenska i år 6 (samma år). Hon jämför läsförståelsedelen i det nationella provet med DLS läsförståelsetest och kommer fram till att DLS-testet ställer högre krav på läsförståelse än det nationella provet gör. Detta resultat antyder att DLS-screeningen i årskurs 7 och 8 kan ge viss vägledning för om eleverna kommer att klara det nationella provet i svenska i år 9. Liknande studie gör Levlin (2014) i sin avhandling och pekar på att det fanns ett tydligt samband mellan låga poäng i en läs- och skrivscreening bland 44 elever i årskurs 2 och att inte bli godkänd på nationella proven i svenska i årskurs 3. De elever som misslyckades i år 3, var redan upptäckta i screeningen i årskurs 2, samtidigt som en del elever hade låga poäng på screeningen och ändå klarade nationella proven. Levlin menar att det är möjligt att använda en screening av läsförmågan i år 2 för att identifiera vilka elever som riskerar att inte klara läroplanens mål. På liknande sätt jämförde Rehn & Sedvallson (2012) resultatet för DLS-testet i läsförståelse med resultaten för testen i läsförståelse på de nationella proven i svenska i år 3 för ett antal elever. De fann ett samband, särskilt vad det gällde delprov B (skönlitterär text) och menar att DLS-testet kan användas för att tidigt på skolåret hitta de elever som inte kommer att nå målen när det gäller läsning av skönlitterär text snarare än faktatext. Helleklev (2013) har också visat att resultaten på en läs- och skrivscreening kan förutsäga vilka gymnasieelever som riskerar underkänt (IG) i flera ämnen långt innan betygen sätts, vilket ger screeningstestet en större betydelse än många lärare och skolledare anser att det har. I studien visas ett samband mellan gymnasieelevers resultat på ett diagnostiskt läs- och skrivprov (LS) (Johansson, 2016) i årskurs ett på gymnasiet och deras betyg på vårterminen i årskurs tre, särskilt tydligt var sambandet mellan stavningsprovet och avgångsbetygen.

Även Jacobson (2014) för i handledningen till lästestet "Läskedjor" fram påståendet att ett standardiserat lästest kan ha ett prediktionsvärde i förhållande till betyg (Jacobson, 2014). Korrelationen mellan testet Ordkedjor i årskurs 2 och medelbetyget i årskurs 9 var

medelstarkt. Ännu högre korrelation hade betyget i svenska i årskurs 9 med resultatet på Ordkedjor i årskurs 2, dock med de betygsvärden från den 5-gradiga betygsskalan som förutsätter att betyg är normalfördelade. Hur testet relaterar med dagens betygsskala behöver undersökas, men Jacobson tror att sambandet skulle vara i samma storleksordning (Jacobson, 2014). Rudenius Sandin (citerad i Johansson., 2016) fann också signifikanta samband mellan resultat i klassdiagnoser i läsning och skrivning (LS) och elevernas samtliga betyg i år 8, förutom idrott och slöjd. På liknande sätt kommer Aldrin m.fl. (citerad i Johansson, 2016) fram till att det finns ett starkt samband mellan LS diagnoser och betyg i svenska i skolor 9.

Det råder konsensus om att läs- och skrivscreeningar bör göras och att resultaten från dessa tester inte alltid tas tillvara på ett sätt som gagnar elevers lärande. Olika typer av studier kommer fram till att resultat från läs- och skrivscreeningar samvarierar med elevers betyg, även med flera års mellanrum.

2.3 Forskning om formativ bedömning

Om svensk forskning om läs- och skriv screeningar och DLS är knapp så är forskningen om formativ bedömning, både i svensk och internationell kontext, desto rikligare. Här ger jag först en översikt över den formativa begreppsutvecklingen för att därefter gå in på skillnaden mellan formativa och summativa bedömningar. Avslutningsvis introducerar jag forskning om formativ bedömning ur en policysynpunkt.

2.3.1 Formativ begreppsutveckling

I forskningsöversikten *Assessment and Classroom learning* från 1998 lägger Black & Wiliam grunden för den teoribildning och de empiriska studier omkring begreppet formativ bedömning som publicerats sedan dess och som främst härrör från den brittiska, australiensiska och nyazeeländska kontexten (Hirsh & Lindberg, 2015). I artikeln *Developing the theory of formative assessment* (Black & Wiliam, 2009) försöker författarna definiera vad formativ bedömning och formativa praktiker är, utifrån teorier om lärande.

Fem nyckelstrategier för formativ bedömning har arbetats fram av Black och Wiliam tillsammans med klassrumslärare, i vilka både lärare och elev och eleverna sinsemellan är aktiva och som ska gagna lärandet genom att göra lärandemålen tydliga för eleven (Black m.fl., 2003). Dessa nyckelstrategier berör bland annat att dela framgångskriterier med

eleverna, kommentera elevprestationer, självbedömning och hur summativa prov kan användas formativt.

Dessa strategier kompletterades senare med tre nyckelprocesser som Wiliam & Thompson (2007) utvecklade från Ramaprasad (1983). Dessa består i:

- Ta reda på var eleven befinner sig i sin läroprocess
- Konstatera vart eleven ska
- Formulera en plan över vad som behöver göras för att komma till målet

Ramaprasad (1983) ansåg att feedback ska fylla kunskapsgapet mellan individens nuvarande nivå och den nivå som eftersträvas. Ramaprasad betonar att enbart informationen om kunskapsgapet i sig inte är feedback. Hattie och Timperley (2007) utvecklade en modell för att minska det så kallade kunskapsgapet och menade att en bra feedback ska innehålla svaret på frågorna: Vart ska jag? (vilka är målen), Hur ska jag komma dit? (vad behöver jag göra eller lära mig? och Vad ska jag göra just nu? (vad blir nästa steg).

Black & Wiliam (2009) vill dock i sin artikel "Developing the theory of formative assessment" skifta fokus från att det är läraren som ska implementera formativ bedömning för eleven i analogi med Wiliam & Thompsons modell (2007), mot att eleverna enskilt och tillsammans ska axla ett ansvar för att lärande sker i klassrummet och utvecklar ytterligare strategier med fokus på elevaktivitet.

Forskningen om formativ bedömning bedöms stå på två ben anser Hirsh (2017), vilket framkom i forskningsöversikten om formativ bedömning av Hirsh & Lindberg (2015). Samtidigt är det uppenbart, menar Hirsh (2017) att det ena benet som berör forskning som har med förandet och utvecklandet av elevens lärande väger mycket tyngre än det andra benet som belyser användningen av bedömningsinformation för att omskapa undervisningen eller förändra miljön i vilken lärandet sker. Hirsh (2017) upplever att forskningen sällan berör hur undervisningen och miljön kan göras annorlunda med hänsyn till data från elevprestationer, däremot handlar många studier om hur till exempel feedback eller självbedömning ska förändra eleverna. Detta illustrerar också Per Blomqvists (2018) avhandling om vad svensklärares bedömning av texter leder till när det gäller didaktiska beslut och betyg. Författaren visar på en distinktion mellan lärarnas bedömning av texter och deras didaktiska beslut i klassrummet. En av studiens slutsatser är att formativ bedömning inte bara ska handla

om hur eleverna kan använda lärarens bedömning för att utvecklas, utan att bedömningarna ska forma undervisningen.

Om synen på formativ bedömning tidigare varit färgad av behavioristiska och kognitiva idétraditioner (se kapitel 5), så har det socio-kulturella perspektivet fått större inflytande de senaste 10 åren menar James (2012) och kallar denna ”den tredje generationens förståelse” av formativ bedömning. Enligt detta perspektiv ses bedömning och lärande som framförallt en kulturell och kontextbunden aktivitet och hänsyn behöver tas till sammanhanget, det vill säga att både lärare och elever har med sig förväntningar in i klassrummet som påverkar såväl summativa som formativa bedömningar. Lärandet anses inte handla om överföring av kunskaper utan i högre grad av gemensamt skapande av kunskaper och resultat av formativ bedömning kan variera när kontexten skiftar. James argumenterar för att den sociokulturella praktiken av formativ bedömning, men klargör också svårigheten att kombinera formativ praktik i klassrummet med storskaliga summativa kunskapsmätningar. Detta sociokulturella synsätt bekräfts av Wiliam & Black (2018) som anser att forskarvärlden fortfarande inte kan enas om en enhällig definition av formativ bedömning, men oavsett detta bör formativ praktik alltid förstås i ljuset av den sociala, kulturella och politiska kontext den sker inom.

2.3.2 Skillnaden mellan summativ och formativ bedömning

Distinktionen mellan formativ och summativ bedömning är omdiskuterad. Harlen (2012) menar att man kan se summativ och formativ bedömning som åtskilda eftersom de tydliggör bedömningars olika syften, men att de båda funktionerna ändå befinner sig inom samma bedömningsspektrum. Harlen betonar idén om att samma prov kan användas både i formativt och summativt syfte. Vidare menar Harlen att ett summativt bedömningstillfälle kan användas i ett formativt arbete och att summativa bedömningstillfällen behöver designas på ett sätt som ger lärarna möjlighet att använda provet i formativt syfte, i relation till uppställda mål och kunskapskrav. Black och Wiliam (2009) menar också, liksom Bennet (2011) att test som har en summativ karaktär kan användas formativt och bidra till ett ökat lärande, både för att få elever att förstå vad som förväntas inom ett visst ämne (framgångskriterier) och att elever kan hjälpa varandra med testet som guide för hur de ska studera (Black & Wiliam, 2009).

Bennet argumenterar liksom Hirsh (2017) för att formativ praxis inom forskning behöver fokusera på förmågan att samla in material om elevernas kunskaper, samt dra slutsatser om

detta och planera framåt. Om en länk i denna trestegsprocess brister hämmas eleven i sin kunskapsinhämtning.

Flera forskare inom fältet går emot tanken att formativ bedömning ska få en betydelse av ”upprepad testning”, alltså att lärare använder summativa tester för att undervisa eleverna mot nästa test med målet inställt på det sista stora summativa testet till exempel nationella proven. Klenowski (2009) och Chappuis (2005) menar att detta är en feltolkning av formativ bedömning eftersom man genom att samla elevdata i realiteten har siktet inställt på högre testresultat och poängen på de summativa testerna blir målet i sig. Den instrumentella förståelsen av formativ bedömning, där resultatet från summativa tester används för att styra undervisningen, är bara en annan variant av ”teaching to the test” menar Hume och Coll (2009).

Det finns konsensus på forskningsområdet omkring det faktum att formativt klassrumsarbete är hotat av statliga krav på kvantifierbara mått av kunskap till exempel nationella prov. Frekvent testning gör det svårt att utveckla formativa bedömningspraktiker som då istället blir ett sätt få eleverna att höja sina resultat på externa testet (Hirsh & Lindberg, 2015). Prov är enligt Jönsson (2018) ett osäkert, kontextlöst sätt att bedöma elevers kunskap på, särskilt i ett kriterierelaterat betygssystem där komplexa förmågor ska bedömas och inte bara faktakunskaper så har inget formativt värde för den enskilde eleven. Däremot kan det användas för en så kallad *formativ utvärdering* (Korp, 2011) på gruppnivå där lärare får en översikt över styrkor och behov hos eleverna inför den fortsatta undervisningen. I en avhandling av Boström (2017) fick lärare i matematik kontinuerlig utbildning inom det som Korp benämner som formativ utvärdering och hittade sätt att få information om elevers kunskapsnivå som kunde användas för att utveckla undervisningen. Detta ledde till förbättrade resultat hos eleverna i en klass, men inte i en annan, vilket bland annat bottnade i tidsbrist och överbelastning. Studien belyser att skolor behöver ett seriöst, långsiktigt och omfattande fortbildningsarbete för att implementera formativ bedömning i klassrummet.

2.3.3 Formativ bedömning på policynivå

Black och Wiliam (1998) kunde i sin översikt visa med hjälp av effektmått att formativa insatser skulle ge mycket starka lärandeeffekter. Resultaten var förbluffande höga och hade inflytande på beslutsfattare i hela världen när det gäller att föra in idéer om formativ bedömning i skolan (Hirsh & Lindberg, 2015; Baird m.fl., 2014). Många länder har sedan dess lyft fram sitt arbete med formativ bedömning, men enligt Lundahl m.fl. (2016) är detta enbart retorik. Den formativa bedömningens framfart har skett samtidigt med att storskalig summativ bedömning till exempel nationella prov och storskaliga internationella kunskapsmätningar, har ökat i många länder, inte minst i Sverige menar Lundahl m.fl., (2016). Hirsh & Lindberg (2015) anser att summativa och formativa praktiker behöver kunna fungera sida vid sida. Läraren kan formativt bedöma individen, men vid givna tillfällen ska elevens kunskaper också testas och jämföras mot nationella kunskapsmål (Hirsh & Lindberg, 2015). I Sverige har de nationella proven sedan införandet 1997 ökat i antal och betydelse (SOU 2016:59; Skolverket 2018a) samtidigt som internationella kunskapsmätningar till exempel OECD:s test PISA har fått en ökad relevans (Lundahl m.fl., 2016).

När stora kunskapsmätningar fått stor betydelse för individen har en ny typ av formativa, kommersiella tester som fått betydelse av *benchmark assessment* (en måttstock) för huruvida elever ska klara sitt slutprov växt fram, framförallt i USA (Fleitz, 2011). En kritik har växt fram mot dessa tester eftersom de anses vara standardiserade tester i en annan förpackning och det handlar om "teaching to the test", att använda formativ bedömning för att undervisa för bättre resultat på ett summativt prov. På grund av det stora antalet standardiserade tester på olika nivåer i det amerikanska utbildningssystemet har den formativa bedömningspraktiken svårt att rota sig där (Hirsh & Lindberg, 2015).

När resultat gällande olika pedagogiska tillvägagångssätt presenteras med hjälp av statistiska granskningsmetoder där styrkan av inlärningseffekter anges med effektmått såsom Black och Williams översikt från 1998 och även John Hatties översikt *Visible Learning* (2009) får de en stor genomslagskraft på utbildningspolitiken i olika länder. I Hatties *Visible Learning* framförs inlärningseffekter för strategier som kan kategoriseras som formativ bedömning, till exempel att eleverna ska utveckla självreglering och metakognitiva strategier samt feedback mellan lärare och elev. Det har rests kritik mot dessa metastudier som varande *kontextblinda*,

det vill säga att den kulturella, ekonomiska, sociala eller historiska kontexten inte beaktas när länder förändrar innehållet i läroplaner och skollagar på grundval av effektstudier (Hirsh & Lindberg, 2015). Dessutom visar forskning att metoder och information om formativ bedömning inte räcker utan att skolor, kommuner och länder behöver satsa både pengar och tid för att lärare kan göra dessa formativa metoder till sina egna och skapa effektiva praktiker i klassrummet, likaså som lärarutbildningar behöver arbeta medvetet med bedömningsfrågor (CERI, 2005). Att formativ bedömningspraktik tar tid och kräver en långsiktigt plan kom Boström (2017) fram till i sin avhandling *Formativ bedömning: en enkel match eller en svår utmaning?* Lärarna hade svårt att utveckla den formativa undervisningen på grund av bland annat tidsbrist och bristande stöd från kollegor och skolledare.

I skriften *Formative assessment: improving learning in secondary classrooms* konstateras också att skolan behöver hitta sätt att kombinera formativa och summativa syften och tillvägagångssätt samt försäkra sig att klassrumsbedömning kopplas till systematiskt kvalitetsarbete på skol- och systemnivå, investera i utbildning och ge stöd till lärare och skolledare att bygga starka broar mellan forskning, policy och praktik (CERI, 2005).

3 Teorianknytning

Synen på bedömning har sedan 1970-talet förändrats från en test- och provkultur till en bredare bedömningskultur där lärarnas bedömningsarbete i klassrummet betonas (Gipps, 1994). Samtidigt har en ökad målorientering och decentralisering inom skolan väckt frågan om likvärdighet mellan skolor när det gäller bedömning (Gustafsson m. fl., 2014; Wikström, 2013; Wedman, 2003). Antal prov som kommer "uppifrån" till exempel nationella prov och nationella bedömningsstöd har dessutom ökat i antal och blivit viktigare för elevernas betyg (Skolverket, 2019d), liksom läs- och skrivscreeningar har ökat i omfattning på landets skolor (Wikström, 2013; Myrberg & Lange, 2006; Johansson, 2016). För en ökad likvärdighet arbetar Skolverket aktivt för att lärarkårens bedömningskompetens ska utvecklas (Skolverket 2019a). När proven, testerna och bedömningsstöden ökar i antal och betydelse behöver skolor veta varför de ger olika typer av prov och hur resultaten ska användas (Wikström, 2013).

3.1 Teoretisk ram

Kunskap om den teoretiska grunden för formativ bedömning är angelägen för alla som arbetar med undervisning. Utan den kan formativ bedömning begränsas till en uppsättning tekniker som utförs mekaniskt i klassrummet. Undervisnings- och bedömningsmetoder bör existera tillsammans med lärandeteorier och pedagogisk filosofi (Pryor och Crossouard, 2008; Crossouard & Pryor, 2012; Hassan, 2011). Lärare bör vid bedömningssituationer vara medvetna om vilken eller vilka lärandeteorier som står bakom de metoder som används (James, 2006; Shepard, 2001).

Tolkningen av formativ bedömning skiljer sig åt beroende på vilken teori om lärande och kunskap som vi utgår ifrån. De tre teoretiska perspektiv som här presenteras i relation till formativ bedömning är psykometrisk/behavioristisk teori, kognitiv teori samt socialkonstruktivism/sociokulturell teori. Teorierna diskuteras i ljuset av förändringarna på bedömningens område som skett i den svenska skolan, vilket bidrar att placera min fallstudie i en historisk kontext.

3.1.1 Psykometrisk tradition

En skiljelinje i synen på kunskap går mellan den empiriska och den rationella skolan, där empirismen strävar efter att testa observerbara data och kvantifiera dessa, medan den rationella skolan ser kunskap som något som finns inom människan och som ska "förlösas" av pedagoger. Empirismen slog an tonen för forskning inom många områden under första hälften av 1900-talet bland annat den psykometriska (psykopedagogisk mätningsslära) traditionen i form av normrelaterad bedömning (Jönsson, 2017). Psykometrin fick stort inflytande i Sverige och USA under 1940-talet då frågan om urvalssystem till fortsatt och högre utbildning började diskuteras allt mer frekvent (Wedman, 2003). Psykometrin påverkades av psykologins intresse för intelligensmätningar som tagit fart i slutet av 1800-talet, där individers mentala förmågor och funktioner objektivt skulle mätas, beskrivas och jämföras statistisk (Jönsson, 2017). Föreställningen om att befolkningens kognitiva förmågor skiljde sig åt på ett så stabilt sätt att man kan testa och ordna individerna i en normalfördelad skala var stark (Wedman, 2003). Intelligenstest skulle göra skolans undervisning mer effektiv då elever direkt skulle kunna placeras på rätt utbildningsnivå och elever i behov av stöd skulle kunna få det (Lundahl m.fl., 2016). Snabbt och effektivt kunde man med hjälp av tester kartlägga många människor och placera dem på rätt plats till exempel inom armén eller skolan. Edward Lee Thorndike (1874–1949) fortsatte i USA att utveckla intelligenstesterna för att undersöka elevers inlärningskapacitet psykometrin kom in i skolan, bland annat för att öka likvärdigheten och göra bedömningen av eleverna mer pålitliga och objektiva (Hirsh & Lindberg, 2015). Psykometrin har haft ett stort inflytande på skolan och blivit norm för hur man bedömer elevers prestationer med hjälp av prov (Jönsson, 2016). Thorndikes teorier om lärande och beteende lade grunden för behaviorismen och påverkade Burrhus Frederic Skinner (1904–1990) som influerat synen på lärande och bedömning utifrån ett behavioristiskt perspektiv (Crossouard & Pryor, 2012). James (2008) argumenterar för att kontrollerade testmiljöer och flervalsfrågor ska ses som en del av den beteendevetenskapliga teoritraditionen med fokus på resultat och inte av tankeprocessen. Kontinuerlig mätning med förändring av undervisningsupplägget i kombination med förstärkning anses också höra till den behavioristiska bedömningstraditionen, inte så långt ifrån idén med frekventa återkopplingar inom formativ bedömning menar Lundahl m.fl. (2016). Inläringen uppdelad i små steg, anpassade efter elevens nivå, som förstärks och behärskas innan eleven går vidare på nästa nivå, vilket minskar det aversiva i inläringen. Cooper, Heron, & Heward, (2007)

beskriver hur B.F Skinner lade grunden till teorin om operant betingning som förklarar lärande och allt mänskligt beteende som en följd av tidigare konsekvenser av beteendet, även tankar och känslor räknas som mänskligt beteende, vilket skiljde sig markant mot tidigare psykologiska teorier (Descartes, Freud och Piaget) som istället bygger på att det existerar hypotetiska och mentala inre dimensioner som styr våra handlingar, men som varken kan observeras eller styras (Cooper m.fl., 2007).

Målet för forskarna inom psykometrin var att utveckla likvärdiga, objektiva tester för att mäta kognition och andra egenskaper, och fortfarande är psykometrisk teori grunden till internationella kunskapsmätningar som PISA, för de svenska nationella proven och för Högskoleprovet (Lundahl m.fl., 2016). Baird m.fl. (2017) diskuterar dock svårigheten med att relatera en teori om inlärning till internationella storskaliga mätningar, eftersom de i stor utsträckning påverkar och påverkas av policy, läroplaner och provdesign. Provkonstruktörer har främst varit intresserad av de tekniska aspekterna av testerna, snarare än hur de påverkar inlärningen och Baird m.fl. (2017) menar liksom Resnick och Schantz (2017) att det nu är dags för psykometrin och den pedagogiska forskningen att samarbeta för att konstruera tester som kan säga något kunskaper som anses viktiga i dagens samhälle.

Psykometrins tankar om norm/grupprelaterad bedömning lade grunden för den svenska grundskolans betygssystem, som framförallt infördes för att ska jämförbart urvalssystem till högre studier (Wedman, 2003). Kunskapsmätningar inom detta system görs genom att varje individs prestation jämförs med gruppens prestationer, liknande DLS-proven. En elevs individuella betyg är beroende av hur gruppen presterar som helhet, inte av individens prestation som sådan. Betygen skulle korrigeras med resultat från centrala prov som genomfördes i hela landet (Wedman, 2003). Dessa provuppgifter konstruerades och prövades ut för att normalfördelningskurvan skulle uppstå. När gruppen presterar bättre höjs kraven och det krävs fler poäng för höga betyg (Hirsh & Lindberg, 2015). Likvärdigheten begränsades dock av att en del lärare satte betygen i relation till hur klassen presterade, alltså klassen blev norm samt att centrala prov saknades i flera ämnen (Hirsh & Lindberg, 2015; Wikström, 2013). Det normrelaterade betygssystemet satte djupa spår i hur man tänker och arbetar med bedömning i svensk skola och har givit prov och summativ bedömning en stark ställning (Jönsson & Balan, 2014).

Betygssystemet i Sverige övergick 1994 från ett normbaserat till ett målrelaterat, vilket var en mycket betydande förändring (Djuvfeldt & Wedman, 2007). Kritiken mot det normbaserade testerna har varit att eleven bedöms efter hur andra presterar, och inte efter hur mycket hon kan och att man förutsätter att det ska finnas en skillnad mellan individer redan från början (Jönsson, 2016). Enligt Roos och Hamilton (2005) relateras psykometri framförallt till summativa bedömningar och behavioristiska grundantaganden. Shepard (2001) för fram, i *Handbook for teaching and learning*, att det psykometriska/behavioristiska sättet att se på mätning av kunskaper lever kvar i hur vi tänker om skolan och kunskap trots att andra teorier om inläring nu påverkar våra aktuella läroplaner. Vissa forskare till exempel Jönsson (2016, 2018) och Lundhal (2014) kritiserar det faktum att prov fortfarande är ett vanligt sätt att mäta kunskap på, och anser att vi helt bör lägga den psykometriska kunskapssynen åt sidan och använda oss av andra bedömningsformer som passar bättre med det nuvarande mål- och kriterierelaterade betygssystemet och med vår syn på kunskap och utbildning.

3.1.2 Kognitivism

Mål-/kriterierelaterad bedömning som infördes 1994 i svenska skolor, bygger på andra grundantaganden om kunskap och lärande, som i sin tur påverkat synen på undervisning och bedömning. Bakgrunden till denna bedömningstradition står att finna i två teoretiska utgångspunkter, som till viss del står i motsättning till varandra; den rationella/undervisningsteknologiska/behavioristiska samt den sociokulturella och sociokonstruktivistiska (Hirsh & Lindberg, 2015).

Det första kan anses representera teknisk rationalitet/undervisningsteknologi med kognitivistiska och även behavioristiska perspektiv och sammanföll i Sverige med en tid (1950–1970) av ekonomisk tillväxt, teknikutveckling, demokratisering, framtidstro och social ingenjörskonst (Wallin 2005). Inom undervisningsteknologin strävades efter effektiv inläring med exakta och tydliga mål som skulle förverkliga varje individs potential genom en systematisk och programmerad inläring där läraren utförde ett program designat av experter (Wallin, 1968). Skinner (1968) har i sin bok *Undervisningsteknologi* beskrivit hur undervisningsteknologin skulle kunna gestalta sig i form av undervisningsmaskiner som skulle ge kontinuerlig återkoppling och förstärkning till studenten och dela upp lärandet i små steg som skulle gås igenom linjärt.

Kognitiva teorier fick stort inflytande på denna syn av bedömning. Inom kognitivismen står individens tänkande i fokus och kunskap ses som något som kan brytas ner i mindre delar och som människan ska ta in utifrån (Hirsh & Lindberg, 2015). Blooms (1956) taxonomi kom att bli tongivande inom denna strömning, med föreställningen att kunskapsinhämtningen sker i en speciell ordning, från kunskaper på låg kognitiv nivå (t ex. fakta) till kunskaper på högre nivå (t ex. förståelse, analys, slutsatser och eget skapande). Särskilt i utformningen av det första kriterierelaterade betygssystemet konkretiserades denna syn på kunskap genom de fyra F:n som stod för fakta, förståelse, färdighet och förtrogenhet (Jönsson, 2018). Wedman och Wallin exemplifierar hur denna inriktning tog sig uttryck i Sverige då Skolöverstyrelsen drev det så kallade MUT (Mål och utvärdering) -projektet mellan 1972 och 1975. Syftet var att i detalj beskriva alla skolämnens mål på olika nivåer och detta gjordes i behavioristisk anda med konkreta beteendevetenskapliga mål för varje ämne. Projektet lades ner efter en diskussion i riksdagen, eftersom man ansåg att lärarnas frihet att utforma undervisningen var hotad (Wedman, 2003; Wallin, 2005). Detta synsätt framstår som ett alternativ till normrelaterad bedömning och de kognitiva kunskapsnivåerna ansågs vara ett uttryck för individens utveckling (Hirsh & Lindberg, 2015). Den formativa bedömningen handlade vid den här tiden (1960- och 70-talen) främst om diagnostiska prov och nivågruppering (Black, 2001). Baird m.fl. (2017) framhåller den kognitiva teorin som det paradigm som idag är närmast kopplat till storskaliga tester, men kommer samtidigt fram till att teorier om inlärning inte är prioritet vid utformningen och analysen av storskaliga tester.

Denna kognitivistiska/undervisningsteknologiska uttolkningen av formativ bedömning genomfördes aldrig i Sverige i stor skala trots att en sådan förändring förbereddes. Förutom att lärarnas frihet ansågs hotad fanns bakgrunden till det minskade intresset i en 68-rörelse som ifrågasatte makten och misstrodde kontroll. Istället för att betona eleven som objekt för undervisning sågs nu individen som ett subjekt i relation med demokratin och inte som en förlängd arm till staten genom snäva beskrivningar av utbildningsmål (Wallin, 2005). Det normrelaterade systemet levde i princip kvar trots att bedömningssystemet hade förändrats och en ny syn på lärande och kunskap dominerade (Selghed, 2004)

3.1.3 Socialkonstruktivistiska/sociokulturella perspektiv.

Den tolkning av målrelaterad bedömning som fick inflytande när Sverige förändrade sitt betygssystem på 1990-talet grundades istället på en kritik av det kognitivistiska synsättet och undervisningsteknologin. Wedman (2003) hävdar att den forskning som fanns om mål och kriterierelaterad bedömning, initierad av bland annat Glaser (1963) och den tongivande artikeln *Instructional technology and the measurement of learning outcomes: Some questions*, ignorerades helt när den nya Läroplan för det obligatoriska skolväsendet, förskoleklassen och fritidshemmet Lpo 94 skrevs. Målen och betygskriterierna skulle istället vara öppna och tolkningsbara för att inte begränsa läraren, utan kunna anpassas till lärarens specifika kunskaper och skolans lokala resurser. Elevernas prestationer och kunskapsutveckling skulle jämföras med målen och kriterierna i läroplanen och betygssättas på hur väl prestationen svarar mot målen (Wedman, 2003). Idén om normfördelning av betyg förkastades och istället skulle varje individs kunnande jämföras med samhällseliga förväntningar uttryckta i mål och kriterier, något som kan anses ha både kognitivistisk som individkonstruktivistisk grund (Hirsh & Lindberg, 2015).

De konstruktivistiska idéerna betonar elevens roll för att få en förståelse för det som lärs ut (Baird m fl., 2014) och ser inte kunskap som något som kan föras över mellan människor. Termen ”konstruktivistisk” syftar på att all mänsklig kunskap är konstruerad och skapas utifrån människans personliga erfarenheter i hennes kontakt med omvärlden. Vissa menar att individen skapar kunskap i kontakt med omvärlden (individkonstruktivism), medan socialkonstruktivism ser lärandet som något som sker i ett socialt sammanhang. Jönsson (2016) menar att en konsekvens av det konstruktivistiska synsättet är att kunskap inte kan mätas eller undersökas på ett objektivet sätt, utan måste tolkas i sin kontext. Med detta sättet att se framstår psykometrin som dekontextualiserad eftersom mätningarna inte är knutna till någon speciell situation utan att kunskapen ska vara generell. Dessutom framstår den psykometriska traditionen som oförenlig med konstruktivism eftersom man bryter ner komplexa kunskaper och procedurer i delar och väljer ut ett stickprov av kunskap vid provkonstruktion (Wikström, 2013).

Synen på kunskap skiljer sig åt mellan kognitivism och individkonstruktivism. Medan de kognitivistiska perspektiven förlägger kunskapen utanför människor, anser de konstruktivistiska perspektiven att kunskap är något som skapas mellan individen och

omvärlden, antingen individuellt eller i samspel med andra (Hirsh & Lindberg, 2015). Inom pedagogiken brukar det socialkonstruktivistiska perspektivet beskrivas som ett sociokulturellt synsätt, som betonar det sammanhang och den miljö som eleven befinner sig i, som en grund för förståelsen och inläringen (Baird m.fl., 2014). Inläringen är ett skapande av mening där både elev och lärare måste vara involverade, som också i hög grad påverkas av våra tidigare erfarenheter (Sälgö, 2000).

Caroline Gipps (1999) är bland många en representant för det sociokulturella synsättet och lyfter i sin artikel *Socio-cultural aspects of assessment* fram bedömning av kunskap som social praktik, alltså att bedömning handlar om de som bedömer och de som bedöms och deras inbördes relation, men att även politiska beslut samt tidsmässiga, sociala och kulturella sammanhang påverkar hur kunskapsbedömningen går till. För att bedömningsprocessen ska bli mer jämlik behöver elever föras in i den processen, och utveckla kompetenser i självutvärdering och reflektion över sitt eget lärande och då är till exempel elevens förståelse för betygskriterier och mål i olika ämnen avgörande (Gipps, 1999).

Hirsh (2017) ser att dagens svenska skolsystem positionerar sig tydligt i det som James (2012) kallar den tredje generationens förståelse av formativ bedömning, grundad i sociokulturell teori där lärandet ses som något som sker i samspel mellan elever och lärare i en tillrättalagd miljö. Hirsh efterfrågar en förståelse för bedömningars formativa tillämpning relaterat till utformningen av undervisning och inte bara till formandet av elever, som enligt Hirsh dominerar forskningen om formativ bedömning.

Trots vårt målrelaterade och socialkonstruktivistiskt inspirerade betygssystem lever en rest av vårt tidigare relativa betygssystem kvar i de nationella proven, som fungerar utjämnande för att stärka likvärdigheten (Wedman, 2003).

3.1.4 Teoretisk utgångspunkt

Det sociokulturella perspektivet (James, 2012) kommer att vägleda tolkningen av resultaten i denna fallstudie. DLS-proven ska förstås i ljuset av den sociala, kulturella och politiska kontext den sker i (William & Black, 2018). Synen på lärande, kunskap och undervisning i den svenska skolan är grundad i sociokulturell teori vilket till exempel tar sig uttryck i att undervisningen ska anpassas till elevers olika behov och lärandet ska ske i samspelet och i

miljön (Hirsh, 2017). Samtidigt upplevs de psykometriska och kognitiva kunskapsmätningarna som effektiva och tillförlitliga i vissa sammanhang (Jönsson, 2018; Järpsten & Taube, 2018; Wedman, 2003). Utmaningen blir då att kunna använda summativa kunskapsmätningarna på ett formativt sätt, till exempel för att förändra undervisningen eller som feedback (Korp, 2011; Black et al, 2003; CERI 2005; Hirsh, 2017).

4 Metod

I detta kapitel beskrivs metoden som jag har använt för datainsamling, en fallstudie av kvantitativ karaktär. Jag beskriver här också urval, tillvägagångssätt, etiska principer och metodkritik.

4.1 Studiens metod

Detta är en fallstudie (en s. k case study) som bygger på data från en utvald skola med särskilda karakteristiska, till exempel att det är en friskola med upptagningsområde från hela regionen samt att undervisningen sker på två språk (svenska och engelska). Enligt (Ejvegård, 2009) är syftet med en fallstudie att ta upp en liten del i ett stort förlopp, och med hjälp av fallet beskriva verkligheten och ge läsaren en uppfattning om verkligheten genom fallet, utan krav på generaliserbarhet. Slutsatserna i en fallstudie kan bara ses som indicier och inte anses representativa. Studien innehåller även kvantifierbara data som ska presenteras i tabeller och diagram, därför är min fallstudie av kvantitativ karaktär (Ejvegård, 2009). Enligt Yin (2007) kan en fallstudie vara en empirisk undersökning där till exempel en företeelse studeras i sin speciella kontext. Utifrån data söker studien svar på frågeställningarna, inte för att kunna generalisera, utan för att utifrån detta specifika fall undersöka resultaten och belysa problemområdet.

4.2 Urval

Ett strategiskt bekvämlighetsurval (Stukát, 2011) har gjorts när det gäller materialet. Anledningen är att elevernas betyg och poäng på delproven för DLS finns tillgängligt på skolan där jag är specialpedagog. Totalt finns 379 elever i årskurs 7 och 8 fördelade på 6 klasser per årskurs. 189 elever är födda 2005, 177 är födda 2004, 4 elever är födda 2003 och 9 elever är födda 2006. Majoriteten av eleverna är mellan 12 och 14 år vid provtillfället i augusti och september 2018. Vid testet i februari kan åldrarna variera mellan 12 och 15 år. Deltagarantalet skiftade något för varje test, på grund av att testen utfördes vid olika tillfällen. Bortfallet för varje prov låg mellan 22 och 43 elever per prov på grund av frånvaro samt i vissa fall problem med de digitala testerna. I några få fall har eleverna inte tagits med i analysen av resultat och betyg, eftersom betyg från tidigare skola saknats. Här presenteras en

tabell över antal elever som gjort varje prov, antalet pojkar och flickor per prov samt hur stor andel av eleverna som gjort proven som läser svenska eller svenska som andra språk.

	Antal elever/prov	Flickor	Pojkar	Svenska	Svenska som andra språk
Rättstavning	341	171	170	254	87
Läshastighet	345	174	171	256	89
Ordförståelse	336	172	164	253	83
Läsförståelse	357	182	175	266	91

Tabell 1. Antalet elever, antal pojkar och flickor, antal elever som läser svenska eller svenska som andra språk per prov.

Det var första gången DLS-proven gjordes i så stor skala på skolan. Det ansågs vara av större vikt att eleverna i årskurs 7 och 8 gjorde DLS proven och man valde därför att inte göra proven med eleverna i årskurs 9. Nästa år kommer proven att göras med alla tre årskurser.

4.3 Tillvägagångssätt

Elevernas resultat på DLS läshastighet och DLS ordkunskap hämtas från den digitala tjänsten Hogrefe Online (Hogrefe testsystem, 2019) där eleverna utförde de två digitala testerna i slutet av augusti 2018. Tillsammans med resultaten från DLS rättstavning och DLS läsförståelse (som genomfördes på papper) fördes elevernas Stanine-värden in i ett Exceldokument där samtliga elever i årskurs 7 och 8 på skolan fanns registrerade klassvis. Elevernas betyg från december 2018 hämtades från den digitala lärplattform där också uppgifter om eleven läste svenska eller svenska som andra språk fanns. Elevernas betyg gavs värdena 1–6 i Excel dokumentet där 1 motsvarar betyget F och 6 betyget A. Ytterligare variabler som fördes in i dokumentet var om eleven läste svenska eller svenska som andra språk samt huruvida eleven hade fått specialpedagogisk hjälp under läsåret 2018/2019. I de samhällsorienterande ämnena visade det sig svårt att få ett betyg i samma ämne eftersom eleverna läser olika samhällsorienterande ämnen på höstterminen i årskurs 7 och 8. På grund av detta valdes betyg i ämnet historia för årskurs 7 och betyg i ämnet samhällskunskap för årskurs 8. Det var angeläget att ha samhällsorienterande ämnen som variabel, då undervisning

i dessa ämnen sker på svenska, medan matematik undervisas och testas på engelska med svensk lärobok.

Information om vilka elever som är under utredning eller som har särskilt stöd samlades in. Eleverna anonymiserades och informationen i Excel dokumentet fördes in i statistikprogrammet Mathematica där resultaten från proven gjordes om till stapeldiagram och låddiagram som visuellt åskådliggör utfallet. Sambandens styrka uppmättes med Pearsons korrelationskoefficient (r) som också brukar kallas korrelationen och som är ett mått på hur perfekt sambandet är (Elbro & Poulsen, 2016). För att åskådliggöra sambanden i datamaterialet valdes dock korrelationskoefficienten bort för att istället presentera resultaten med hjälp av låddiagram.

4.3.1 Studiens trovärdighet

Bryman (2011) menar att det finns fördelar med att använda sekundärdata; alltså data som redan har samlats in. Det kan bland annat vara mer ekonomiskt ur ett tids- och pengaperspektiv. Samtidigt kan sekundärdata misstolkas, bland annat därför att avståndet från primärkällan är större. I min studie finns närhet till primärkällan då jag var närvarande när eleverna som genomförde testerna. Jag har också delvis varit ansvarig för rättning av materialet. Studiens trovärdighet bygger framförallt på DLS-provens mätsäkerhet (reliabilitet) och validitet som nu ska diskuteras.

Reliabilitet

Hur vi tolkar resultaten från ett prov beror i hög grad på provets reliabilitet eller tillförlitlighet (Wikström, 2017). Reliabiliteten kan uppskattas med olika metoder som anger hur stort mätfel ett prov har och som visar en elevs "sanna" poäng (poängen eleven skulle få om denne skulle testas ett oändligt antal gånger med samma prov). Reliabiliteten i proven rättstavning och ordförståelse har beräknats med Cronbach's alpha koefficient. Alphakoefficienten kan variera mellan 0 och 1, där 0 innebär att provet inte alls har någon precision och 1 att det exakt mätsäkerhet. Reliabilitetskoefficienten för proven Rättstavning och Ordförståelse är mellan 0,70 och 0,90, vilket kan ses som fullt acceptabelt. Provet läshastighet är mycket tidsbegränsat och detta gör att alla elever inte hinner göra alla testuppgifter och det finns färre svar på uppgifterna i slutet av testet. Istället för Cronbach's alpha beräknas testets interna konsistens istället genom att uppgifterna delades upp i två delar där hälften av

provuppgifterna var udda och den andra hälften jämna och de korrelerades med varandra för att undersöka i vilken utsträckning de överensstämde med varandra. Korrelationskoefficienten visade sig ligga mellan 0,97 och 0,98 vilket visar på en god samstämmighet och tillfredsställande intern konsistens (Järpsten & Taube, 2018).

Roe (2015), Fridolfsson (2015) och Järpsten och Taube (2015) anser att ett hot mot provens tillförlitlighet är att lärarna inte följer anvisningarna för proven som finns beskrivna i lärarhandledningen. I detta fall var alla testledare informerade om hur provet skulle genomföras och alla som genomförde proven hade försetts med skriftligt instruktioner. Vid de digitala testerna ges instruktionerna dels muntligt av läraren, skriftligt på tavlan och på datorskärmen. Trots detta blev det vissa svårigheter vid genomförandet för några elever på grund av det digitala provet.

Mätsäkerheten kan även påverkas av att elever "har en dålig eller bra dag" eller är väldigt nervösa. Dessa mätfel kan påverka negativt eller positivt på elevernas resultat och tar därför ut varandra, menar Jönsson (2018). På gruppnivå får inte detta inflytande, men däremot kan det göra det på individnivå. Ett prov med många och relativt korta uppgifter, som DLS-proven, har högre tillförlitlighet än ett prov med få och längre uppgifter (Jönsson, 2018). Det faktum att proven kommer göras om i augusti och kan studeras om igen ökar också reliabiliteten i materialet.

Enligt Järpsten & Taube (2018) finns det också en signifikant interkorrelation mellan de tre proven rättstavning, läshastighet och ordförståelse som standardiserades på samma gång. Detta betyder att proven mäter liknande begrepp och en elev som presterar ett högt resultat på ett av proven också uppvisar höga resultat på de andra två proven.

Validitet

Validiteten handlar om provet mäter det som är tänkt att det ska mäta (Wikström, 2017).

Ett sätt att mäta validiteten är att studera hur testet samvarierar med andra närliggande test, begrepp eller tecken på det kriterium som testet avser att mäta, så kallad *samtidig validitet* enligt Patel och Davidson (2015). Validiteten bedömdes genom att lärare i svenska innan testningen fick bedöma eleverna på en tregradig skala: svagpresterande elever (20 % av eleverna), medelpresterande (60 % av eleverna) och högpresterande (20 % av eleverna) som sedan korrelerades mot elevernas faktiska resultat på provet. Det visade sig samvariera i hög grad; lärarna kunde förutsäga eleverna resultat korrekt (Järpsten & Taube, 2018). I den tidigare standardiseringen som gjordes 2002 hade de lärarna gjort en korrekt bedömning i 36

% av fallen och de flesta felbedömningarna handlade om att svagpresterande elever hade felaktigt blivit bedömda som medelpresterande (Järpsten, 2002). DLS-testernas validitet prövades också (2002) då elever i årskurs 7–9 testades med testet ”Läskedjor” (Jacobson, 2014) och läsförståelseprovet II ur ”Klassdiagnoser i läsning och skrivning – LS” (Johansson, 2016). Resultaten för eleverna på testerna korrelerades mot deras resultat på DLS-proven rättstavning, ordförståelse, läshastighet och läsförståelse. Korrelationerna fanns signifikanta, särskilt tydligt var detta i testen Meningskedjor (0,37–0,51) och Läsförståelse II (0,42–0,56) (Järpsten, 2002). Den så kallade *innehållsvaliditeten* beskrivs under avsnittet om DLS-proven i kapitel 1. Innehållsvaliditeten är god när provet har en bra täckning av det område man avser mäta till exempel elevers läs- eller ordförståelse (Patel & Davidson, 2015).

4.3 Etik

De elever som kommer ingå i studien kommer inte informeras om att deras betyg och screeningresultat används i denna studie. Vetenskapsrådet (2017) skriver att detta är vanligt inom bland annat så kallade registerstudier. Sådan forskning kan ändå vara känslig ur integritetssynpunkt och värdet av ny kunskap måste ställas mot risken för integritetsöverträdelse. I uppsatsen kommer materialet helt att avpersonifieras. I rapporten redovisas resultaten enbart på grupp- och inte på individnivå, vilket gör att rapporten inte kommer att bryta mot etiska riktlinjer. Samtidigt kan man utifrån detaljer förstå vilken skola som avses, vilket kan innebära tveksamhet till att publicera materialet offentligt. För att erbjuda ökad diskretion har varken namn, skola, stad eller kommun dokumenterats i texten. Elevernas betyg samt resultaten från diagnostiskt läs- och skrivtest omfattas av sekretess eftersom skolan är en friskola. Detta gäller så länge som eleverna går kvar i skolan, därefter ska friskolan lämna betygsuppgifter till kommunen (SFS 2010:800).

5 Resultat

Här följer en redovisning av resultaten utifrån uppsatsens frågeställningar. Första delen granskar resultaten utifrån den normerande stanineskalen. Här görs jämförelser mellan klasser och mellan elever som läser svenska och svenska som andra språk. Med diagram visas också hur fördelningen av stöd ser ut för de elever som har fått de lägsta resultaten (Stanine-värde 1–3). I den andra delen redogör jag för sambandet mellan elevernas resultat i rättstavningsprovet och deras betyg i tre olika ämnen.

5.1 Elevernas resultat

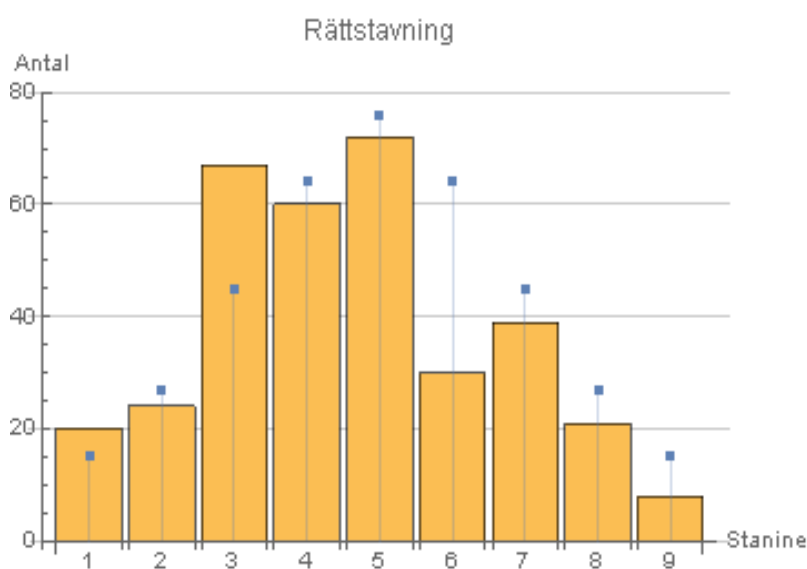
Min första frågeställning handlar om hur eleverna presterade på de olika testerna i förhållande till standardskalan som Järpsten och Taube (2018) har tagit fram och där medelvärde i testerna anses ligga två standardavvikelser från värdet 5 på stanineskalan, mellan Stanine-värde 4–6. Genomsnittligt Stanine-värde på eleverna var 4,6 för rättstavning, 4,2 för läshastighet, 4,5 för ordkunskap och 4,9 för läsförståelse. Således presterar skolans elever på en genomsnittlig nivå. I tabellen nedan redovisas hur elevernas resultat fördelar sig procentuellt på Stanineskalstegen för de fyra olika proven.

Stanine	Procent	Procent	Procent	Procent
	Rättstavning	Läshastighet	Ordkunskap	Läsförståelse
1	5,9	7,8	10,1	5,0
2	7,0	18,0	12,5	9,8
3	19,6	12,8	10,4	8,1
4	17,6	16,1	20,2	12,6
5	21,1	18,6	15,2	16,8
6	8,8	14,5	11,0	19,3
7	11,4	7,0	9,8	18,8
8	6,2	4,0	7,1	6,7
9	2,3	1,4	3,6	3,4

Tabell 2. Fördelning av elevernas Stanine-värden i procent per prov

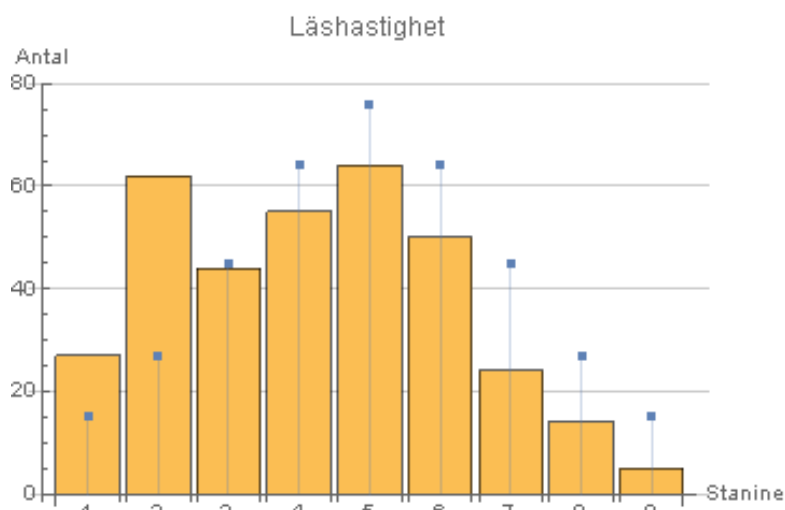
Ungefär lika många flickor som pojkar genomförde varje prov. Vid jämförelse av flickors och pojkars resultat för proven kunde ingen större skillnad observeras, därför åskådliggörs detta inte i stapeldiagram.

Elevernas resultat blir tydligare i följande stapeldiagram över provet rättstavning. På den vågräta axeln visas Stanine-värden från 1–9 och på den lodräta axeln antal elever som fick respektive värde. De blå kvadraterna visar standardskalan som utgör normen för hur eleverna bör prestera.



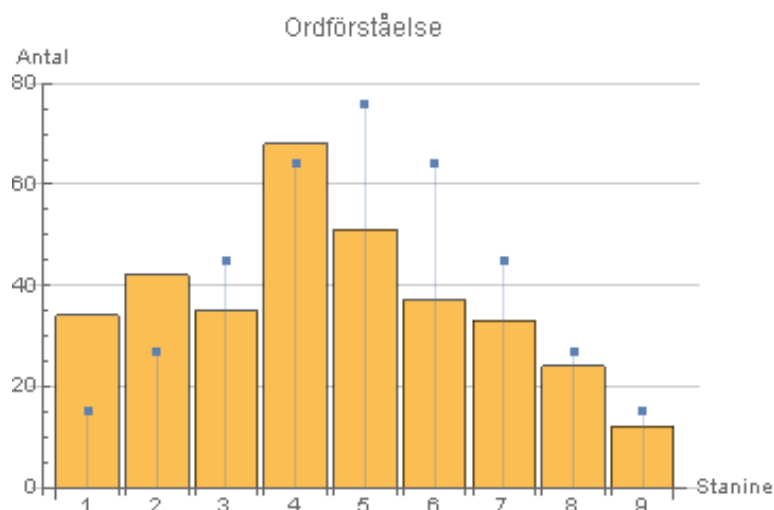
Figur 2. Antal elever per Stanine-värde på rättstavningsprovet

Figur 2 visar att resultatet på rättstavningsprovet ungefär följer den normerande stanineskalan som representeras av de blå kvadraterna. Fler elever fick Stanine-värde 3 än förväntat och dessa elever riskerar, liksom de med Stanine-värde 1 och 2, att ha svårigheter med sin läs- och skrivutveckling och behöver utredas vidare. Färre elever än förväntat finns på skalsteg 6 och över. På läshastighetsprovet illustrerat i figur 3 här nedan kan ett delvis liknande resultat observeras.



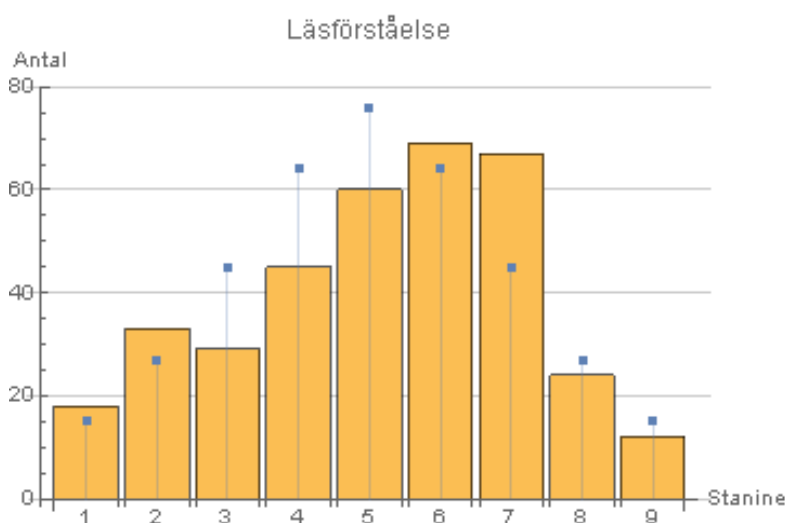
Figur 3. Antal elever per Stanine-värde för läshastighetsprovet

På läshastighetsprovet i figur 3 ser vi fler elever än normen har fått resultat motsvarande Stanine-värde 1 och 2, de lägsta stegen på standardskalan. Färre elever når genomsnittliga och höga Stanine-värden än normen. Liknande förskjutning mot negativa resultat kan observeras på ordförståelseprovet som presenteras nedan, nästan dubbelt så många elever än normen har fått Stanine 1, det allra lägsta resultatet.



Figur 4. Antal elever per Stanine-värde på ordförståelseprovet

En förskjutning åt negativa resultat kan återigen observeras, det vill säga att fler elevers resultat hamnar på Stanine-värde 1-3 än den normerande Stanine skalan. Det är ca 35 elever som får Stanine-värde 1 och strax över 40 elever som får Stanine-värde 2, vilket väcker frågor rörande elevernas ordförståelse. Det prov där eleverna kan sägas ha klarat sig bäst i är läsförståelse presenterat nedan i figur 5.

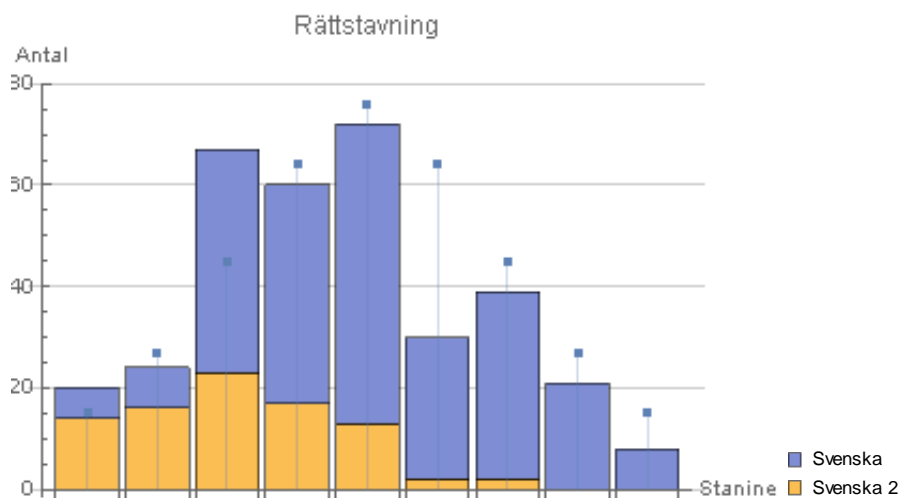


Figur 5. Antal elever per Stanine-värde på läsförståelseprovet

I läsförståelseproven får fler elever än normen Stanine värde 6 och 7 vilket synliggör att eleverna har presterat bättre på detta prov än på övriga DLS-proven. Vi finner att färre elever får poäng som motsvarar de lägsta skalstegen om vi jämför med övriga tester.

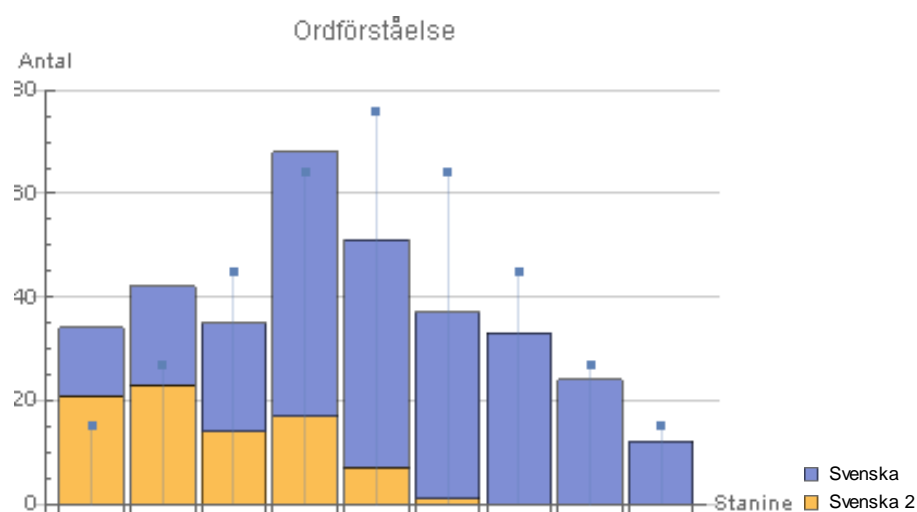
5.1.1 Resultatet för elever som läser svenska och svenska som andra språk

En skillnad noterades mellan elever som läser svenska och svenska som andra språk (sva), vilket visade sig framförallt i rättstavnings- och läshastighetsprovet, men även i de två andra två proven som kan ses i bilaga 1. Resultaten för eleverna som läser sva är gulmarkerade.



Figur 6. Resultatfördelning elever som läser svenska och svenska som andra språk för provet rättstavning.

Elever som läser sva är i majoritet bland de som får de lägsta poängen på proven. På skolan läser elever med svenska och svenska som andra språk tillsammans, får samma undervisning och i stor utsträckning liknande uppgifter, men bedöms olika. Skillnaden mellan de två elevgrupperna är också tydlig i ordkunskapsprovet som följer här nedan.



Figur 7. Resultatfördelning elever som läser svenska och svenska som andraspråk ordförståelseprovet.

I figur 6 och 7 kan vi avläsa att elever som läser sva är överrepresenterade bland de elever vars resultat hamnar inom riskzonen mellan Stanine-värde 1–3 och att det är ovanligt att dessa elever får högre poäng än Stanine 5.

5.1.2 Skillnader klassvis

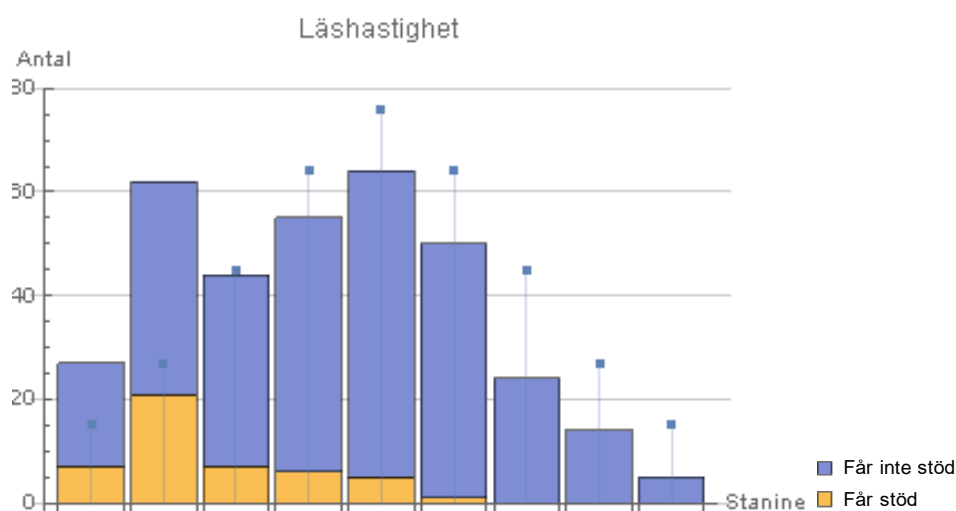
Provresultaten skiljde sig åt mellan olika klasser. I till exempel rättstavningstestet hade den bästa klassen 5,3 och den sämsta 4,0 som genomsnittligt Stanine-värde. Enligt den normerande Stanineskalen från 2018 bör ett genomsnittligt resultat ligga mellan Stanine-värde 4–6. I läshastighetstestet skiljer ett helt skalsteg mellan två klasser (4,7 och 3,7), liksom i ordförståelsetestet (3,9 respektive 4,9). I läsförståelsetestet skiljer sig medelvärdet något mindre mellan klasserna (5,7 och 4,9).

Klass	Medel Rättstavning	Medel Läshastighet	Medel Ordförståelse	Medel Läsförståelse
7A	4,4	3,9	4,5	5,2
7B	4,9	4,2	4,5	5,1
7C	5,3	4,1	4,5	5,0
7D	4,8	3,7	3,9	5,0
7E	4,8	4,1	4,9	5,7
7F	4,3	3,6	4,1	5,1
8A	4,7	4,2	4,5	5,5
8B	4,2	4,7	4,6	5,0
8C	5,1	4,6	4,3	5,0
8D	4,0	4,6	4,2	5,1
8E	4,1	4,5	4,7	5,1
8F	4,6	4,2	4,9	4,9

Tabell 3. Klassernas genomsnittliga Stanine-värden vid respektive prov

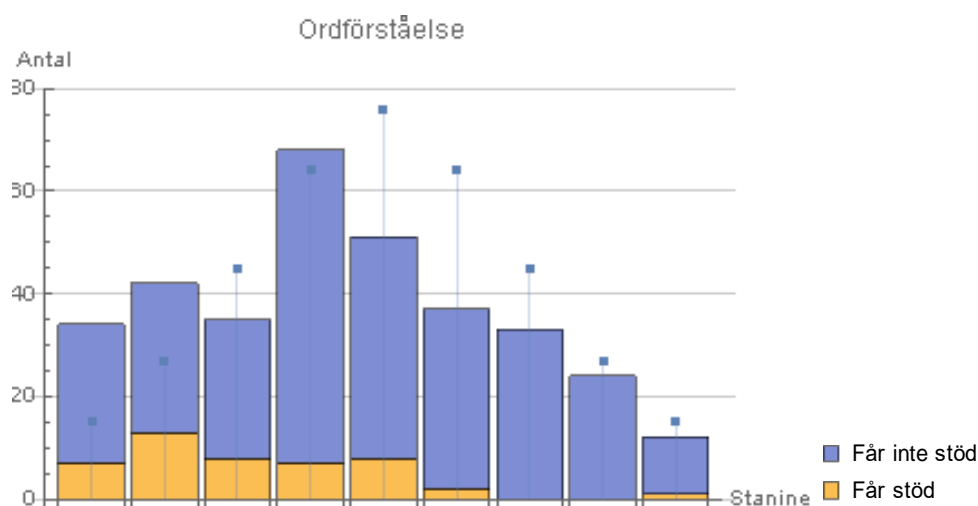
5.1.3 Förhållandet mellan låga resultat på DLS-proven och särskilt stöd

Ett syfte med DLS-proven är att identifiera elever som presterar under genomsnittet med ett Stanine-värde mellan 1–3 och följa upp med tester, intervjuer och eventuellt en pedagogisk utredning samt särskilt stöd. Målet är att eleverna ska få hjälp att komma ifatt i sin läsning och skrivning (Järpsten & Taube, 2018; Roe, 2015; Fridolfsson, 2015). I nedanstående diagram visas hur många elever som har fått eller får särskilt stöd under läsåret 2018/2019, alltså hur många elever som har regelbunden individuell kontakt varje vecka med antingen en specialpedagog, en studiehandledare eller en assistent som antingen genomför en utredning eller hjälper eleven med läsning och skrivandet inom ramen för olika skolämnena. Resultaten visas för proven läshastighet och läsförståelse av utrymmesskäl, men stapeldiagrammen för övriga två prov ser ungefär lika ut på övriga två tester och kan studeras i bilaga 2.



Figur 8. Andel elever som får särskilt stöd relaterat till resultatet i läshastighetstestet

Utifrån figur 8 kan vi avläsa att inte ens hälften av eleverna som får Stanine-värde 1–3 i läshastighetstestet får regelbunden individuell hjälp och har inte heller varit i kontakt med specialpedagog för utredning, vilket är anmärkningsvärt. I figuren här nedan visar hur fördelningen av stöd ser ut för resultaten i ordförståelseprovet som också gjordes i slutet av augusti 2018.



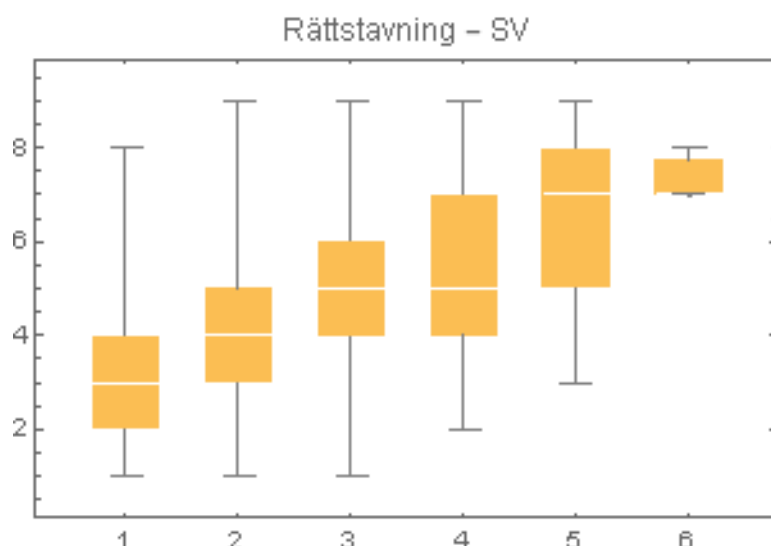
Figur 9. Andel elever som får särskilt stöd relaterat till resultatet i ordförståelseprovet

Figur 9 visar att bara en liten andel av eleverna med de lägsta Stanine-värden 1, 2 och 3 på ordförståelseprovet får specialpedagogisk hjälp. Vi ser också att ett antal elever som presterar genomsnittligt och över genomsnittet på proven får stöd. Det kan handla om elever med svårigheter som inte direkt påverkar läs- och skrivkapaciteten eller att elever som får tillgång till specialpedagogisk faktiskt klarar DLS-proven bättre.

5.1.4 Sambandet mellan resultat på DLS-proven och betyg

Ett centralt syfte med undersökningen var att utreda sambandet mellan DLS-provens resultat och elevernas betyg i svenska/svenska som andra språk, matematik och historia/samhällskunskap. För att illustrera mina resultat använder jag mig av låddiagram som är lämpliga för illustration av spridning i ett statistiskt material. Jag väljer att enbart redovisa resultaten från DLS-provet i rättstavning av utrymmesskäl eftersom resultaten är likartade för övriga tre prov. Övriga låddiagram åskådliggörs i bilaga 3, 4 och 5.

I låddiagrammet i det följande redovisas samvariationen mellan resultaten på rättstavningsprovet och betygen i svenska/svenska som andra språk. På vågrät axel är betygen markerade från F-A. (F=1, E=2, D=3, C=4, B=5, A=6) och på lodrät axel kan de Stanine-värden eleverna fick på provet rättstavning ses (1–9).

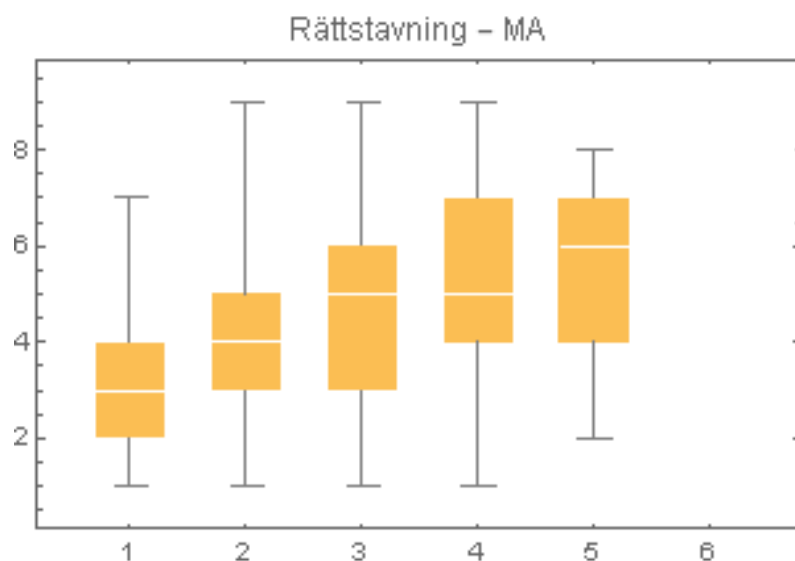


Figur 10. Låddiagram över sambandet mellan elevernas betyg i svenska och resultat på rättstavningsprovet

I den första "lådan" från vänster finns eleverna med betyg F, motsvarande siffra nr 1 på den vågräta axeln. "Lådan" för denna grupp innehåller 50 % av resultaten som ligger mellan Stanine 2–4. Medianen för de Stanine-värden eleverna denna grupp erhållit är 3, den vita linjen på lådan. Underkanten på lådan är den nedre kvartilen (från kvart som betyder en fjärdedel). Från lådans nedre kant och till slutet på utlöparen som slutar med en litet streck befinner sig nästa kvartil med 25 % av elevernas Stanine-värden. Det lilla horisontella strecket indikerar det lägsta Stanine-värde som elever fått vilket är Stanine 1. Det betyder att 75 % av eleverna med F i svenska har Stanine-värden mellan 1-4, medan 25 % av eleverna har fått över Stanine 4. Överkanten på lådan är den övre kvartilen och ovanför den, på den smala utlöparen, finns 25 % av resultaten, varav det högsta värdet är Stanine 8. En elev har alltså fått F i svenska, men ändå Stanine-värde 8 på rättstavningsprovet, vilket är intressant. Detta kan jämföras med till exempel eleverna med betyget B i svenska (siffran 5 på den vågräta axeln) för vilken medianen är 7. Majoriteten i gruppen, 75 %, har ett Stanine-värde mellan 5 och 9 och 25 % av eleverna har ett värde under Stanine 5. Det lägsta Stanine-värde elever med betyget B får är 3 på rättstavningstestet.

För nästan varje betygssteg kan en ökning av medianen för Stanine-värden observeras och majoriteten av eleverna får högre poäng på DLS-proven ju högre betyg de har, vilket indikerar en positiv samvarians mellan betyg och resultat.

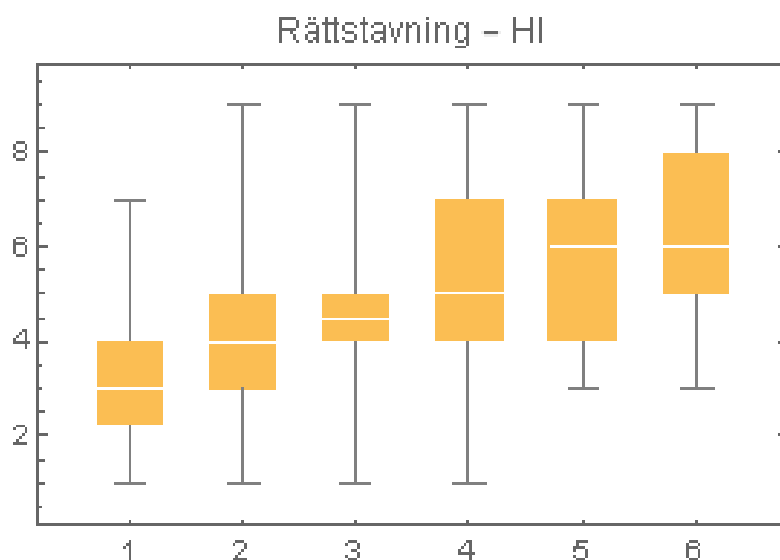
I nästa låddiagram ser vi sambandet mellan resultat på rättstavningsprovet och betyget i matematik. På vågrät axel är betygsstegen presenterade från F-A (F=1, A=6), på lodrät axel har Stanine-värden från 1–9 märkts ut.



Figur 11. Sambandet mellan elevernas betyg i matematik och resultat på rättstavningsprovet

I figuren kan vi avläsa att 50 % av eleverna med betyget F, motsvarande siffran 1 på vågrät axel, i matematik får mellan Stanine-värde 2-4 på DLS-provet rättstavning. 25 % får under Stanine 2 och 25 % får över Stanine 4. Medianen är liksom när det gällde ämnet svenska 3 och maxvärdet som en elev fått är 7.

För betyget B (5 på vågrät axel) i matematik har 50 % av eleverna fått Stanine-värde mellan 4 och 7. En fjärdedel i denna grupp har under 4 medan en fjärdedel har över Stanine 7. Lådan för betyget A saknas eftersom det krävs minst fyra elever för att det ska bli en låda i diagrammet. Medianen för Stanine-värden för betygen B är 6 och för C 5, även om den övre percentilen för både C och B är 7. I nästa låddiagram (Figur 12) presenteras sambandet mellan betyg i historia/samhällskunskap på vågrät axel med betygsvärden F-A (1–6) och resultaten på DLS rättstavningsprov på lodrät axel med Stanine-värde 1-9.



Figur 12. Sambandet mellan elevernas betyg i historia/samhällskunskap och resultat på rättstavningsprovet.

Samma mönster som observerades i figur 10 och 11 kan ses här i figur 12 för eleverna med F i ämnena historia och samhällskunskap. Tre fjärdedelar av eleverna i denna grupp får under Stanine 4. Detta kan jämföras med att tre fjärdedelar av eleverna med A i dessa ämnen ligger över Stanine 5.

Oavsett ämnesbetyg ser relationen till DLS-provens resultat liknande ut. Majoriteten av elever med F eller E i betyg faller under Stanine 4 på rättstavningsprovet och övriga prov (se bilaga 3, 4 och 5). För betyget C, B eller A får majoriteten av eleverna över Stanine 4 på provet. Samtidigt finns det elever som avviker från mönstret i nästan alla betygsgrupper, till exempel har elever med betyget C i historia mellan Stanine 1–3 på rättstavningsprovet, ett resultat i ”riskzonen”, samtidigt som elever med F i historia har fått Stanine-poäng mellan 4 och 7. Dessa elevers enskilda resultat bör följas upp med intervjuer. Den generella bilden som låddiagrammen åskådliggör är att det finns ett samband mellan DLS resultaten från alla fyra proven (se bilaga 3, 4 och 5) och betygen i de tre ämnena som jag undersöker. Lägre resultat samvarierar med lägre betyg och högre resultat samvarierar med högre betyg. Särskilt värden över eller under Stanine 4 är betydande för huruvida elever har F eller högre än C i betyg, vilket överensstämmer med DLS-handledningen (Järpsten & Taube, 2018).

6 Diskussion

6.1 Resultatdiskussion

I detta kapitel kommer resultaten att diskuteras utifrån undersökningens syfte, frågeställningar, tidigare forskning och teori. Därefter följer ett resonemang om metodens för- och nackdelar samt studiens kunskapsbidrag till den specialpedagogiska forskningen. Slutligen ger jag förslag på vidare forskning.

Elevernas resultat och formativ användning av DLS

DLS-proven är ett vanligt screeninginstrument på svenska skolor för att undersöka elevers läs- och skrivförmåga (Järpsten & Taube, 2018; SOU 2016:59). Proven härrör ur en empirisk och psykometrisk tradition med strävan att objektivt mäta och kvantifiera kunskaper för att därefter inordna individer i en normalfördelad skala (Jönsson, 2016; Wedman, 2003; Hirsh & Lindberg, 2015). DLS hör alltså till den psykometriska provtraditionen som också benämns summativ (Jönsson, 2018), men den ökande användningen av provet har snarare att göra med forskningsrön och styrdokument (Roe, 2015; Fridolfsson, 2015; SFS 2010:800) som rekommenderar ett omfattande och kontinuerligt kartläggningsarbete i det förebyggande arbetet med barn och ungdomars läs- och skrivförmåga, än skolors preferens för den psykometriska idétraditionen. Hirsh (2017) menar att den nuvarande läroplanen är grundad i den sociokulturella teoritraditionen där lärandet inte ses som någonting som ska överföras från lärare till elev utan något som skapas i samspelet och i skolmiljön. I inledningen väcker jag tanken om att DLS-resultaten kan vara underutnyttjade i relation till den tid och de resurser som tas i anspråk för genomförandet, något som delvis bekräftas av andra mindre studier i avsnittet om tidigare forskning (Myrberg & Lange, 2006; Bäckman, 2011; Björnsson & Hellström, 2017; Bylund, 2011; Wiktorrell, 2016; Carlsson, 2017). Tanken med proven är att resultaten ska kunna fungera undervisningsjusterande och utvärderande för lärare och resursbeslutsgrundande för skolledare, inte bara som en gallring av elever som har svårare än genomsnittet med sin läsning och skrivning (Järpsten & Taube, 2018; Johansson, 2016; Roe, 2015; Fridolfsson, 2015). Vid formativ bedömning bör hänsyn tas till den sociala, politiska och kulturella kontext den sker i (William & Black, 2018), vilket hjälper oss att tolka resultaten och vägleder de åtgärder som bör tas för att utveckla elevernas lärande utifrån de förutsättningar som finns på den specifika skolan. Att göra testerna bara för att konstatera att

vissa elever presterar sämre än andra är inte förenligt med nuvarande läroplan Lgr11 (Skolverket, 2018) och inte heller med den sociokulturella synen på lärande.

När resultaten från skolan i min fallstudie analyseras, kan det konstateras att trots att skolan är tvåspråkig och åtta ämnen undervisas på engelska, ligger det genomsnittliga Stanine-värdet för resultatet i de fyra proven mellan 4,2 och 4,9 med ingen större skillnad mellan pojkar och flickor. Stapeldiagrammen visar att skolans resultat följer delvis den normerande Stanine-skalan, i vilken genomsnittsliga resultat hamnar mellan 4 och 6. En viss negativ förskjutning mot svagare resultat kan observeras när det gäller proven i läshastighet, ordförståelse och delvis rättstavning. 10 % av eleverna (34 elever) har till exempel Stanine 1 på provet i ordförståelse och 27 % av alla elever som gjorde läshastighetsprovet har Stanine-värde 1–2, vilket kan anses oroväckande. Prestationen på läsförståelseprovet är något bättre och underlättas eventuellt av att rättningen av provet görs enbart på grundval av flervalsfrågor medande skriftliga svaren påverkar inte poängen.

Elevernas svårigheter verkar större med avseende på de tekniska delarna av läsning och skrivning sett till resultaten på rättstavning och läshastighetseproven, vilket kan ha betydelse för läsförståelsen och skrivande enligt modellen *avkodning x förståelse av språk = läsförståelse* eller (Gough & Turner, 1986; Järpsten & Taube, 2018). Samtidigt kan problem med de innehållsmässiga delarna av läsningen observeras i ordförståelseprovet, vilket försvårar läsförståelsen.

Många forskare pekar på den gynnsamma effektanvändningen av formativ bedömning eller utvärdering av summativa prov på gruppnivå kan få på elevernas lärande (Black m.fl., 2003; Black & Wiliam, 2009; Korp, 2011; Harlen, 2011; Bennett, 2011; Black & Wiliam, 1998).

Blomqvist (2018), Hirsh (2017) samt Hirsh och Lindberg, (2015) betonar att formativ bedömning i högre grad bör handla om formandet av undervisningen än formandet av eleven, något som enligt James (2012) ingår i en den så kallade ”tredje generationens” syn på bedömning skapad ur ett sociokonstruktivistiskt och sociokulturellt sätt att se lärande. Att den formativa praktiken bör sträva efter att information från bedömning ska förändra själva undervisningen och undervisningsmiljön hör hemma i en sociokulturell syn på inläring och lärande i vilken lärandet sker i ett samspel och inte enbart från läraren till eleven (Hirsh, 2017; Hirsh & Lindberg, 2015; James, 2012). Utifrån DLS-provens resultat skulle till exempel ett arbete på skol- och gruppnivå omkring elevers ordförståelse kunna initieras då vi sett att 33 % av eleverna i årskurs 7 och 8 hamnar i riskzonen för att ha problem med sin ordförståelse. Järpsten & Taube (2018) rekommenderar DLS som ett utvärderingsmaterial efter en sådan

riktad intervention. En utvärdering av den omformade undervisningen görs när provet genomförs igen tidigt nästa läsår (Järpsten & Taube, 2018). Resultatet kan komma att påverkas något av att eleverna ”lär sig” proven (Jönsson, 2018).

Om datamaterialet från DLS-proven resulterar i en justering av lärandet som medför ett förbättrat elevresultat används proven formativt (Black & Wiliam, 1998; Black m.fl., 2003; Black & Wiliam, 2009; Hirsh & Lindberg, 2015; Wikström, 2013) och kan ses som ett resultat av ett sociokulturellt sätt att se på lärandet (Hirsh, 2017; James, 2012). Samtidigt finns tanken om att använda information från bedömningar för att justera undervisningen både inom den beteendevetenskapliga som inom den kognitiva synen på lärande (Lundhal m.fl., 2016; Cooper m.fl., 2007; Wallin, 2005).

Elever med svenska som andra språk

Elever som läser svenska som andra språk får i högre utsträckning än elever som läser svenska, låga resultat på DLS. Tydligast var detta i testen rättstavning och ordförståelse (figur 6 och 7), men samma tendens sågs i alla fyra proven (se bilaga 1). Det är angeläget att lärare och skolledning har kunskap om att det förhåller sig på detta sätt. Att ta hänsyn till den historiska, ekonomiska, sociala och kulturella kontexten i en skola är en del av den sociokulturella bedömningstraditionen (James, 2012; Black & Wiliam, 2018). Att synliggöra situationen och veta var elever med svenska som andra språk befinner sig i sitt lärande är första steget i den formativa bedömningsprocessen som till exempel Wiliam & Thompson (2007), Hirsh (2017) och Korp (2011) diskuterar. Det finns fördelar att göra detta med DLS eftersom proven har god validitet och reliabilitet (Järpsten & Taube, 2018) och är frivilliga. Följande steg i bedömningsprocessen enligt Wiliam & Thompson är att ställa upp mål och beskriva hur eleven ska ta sig dit. Att på skol- och gruppnivå sätta upp mål och stödja elevers ordförståelsekunskap genom att omforma undervisningen gagnar i synnerhet elever med svenska som andra språk.

Elever i behov av särskilt stöd

Det framkommer att en mycket liten del av de elever som hamnar i riskzonen på DLS-proven (skalsteg 1–3) får vidare utredning av sina svårigheter eller särskilt stöd. Av de totalt 27 elever som hade Stanine 1 på läshastighetsprovet har 7 elever haft eller har regelbunden

kontakt med specialpedagog och liknande andel finns för övriga provresultat. Detta visar på ett behov av att utveckla den formativa bedömningen även på individnivå, följa upp resultaten, ge feedback och sätta upp individuella mål på det sätt som rekommenderas i DLS handledningen (Järpsten & Taube, 2018; Johansson, 2016; Roe, 2015; Fridolfsson, 2015). Situationen påminner om ”testning för testningens skull” (Hagtvét m. fl., 2016; Myrberg & Lange, 2006). Resultaten från DLS underutnyttjas, inte bara på grupp utan också på individnivå på den aktuella skolan, även om några kvalitativa studier kommer fram till att DLS fungerar bra för att hitta och hjälpa enskilda elever på andra skolan (Björnsson & Hellström, 2017; Wiktorell, 2017). Elevens rättigheter och skolans ansvar för elever som inte når kunskapsmålen som beskrivs i skollagen och läroplanen (SFS 2011:800; Skoverket, 2018) är ännu inte en realitet på den aktuella skolan. Dock menar Jönsson (2018) att provresultat enbart bör användas på gruppnivå, eftersom de individuella resultaten alltid är belagda med mätfel. DLS rekommenderas dock att användas formativt på individnivå (Järpsten & Taube, 2018) på det sätt som förs fram som formativ bedömning av Black & Wiliam (1998) samt av Skolforskningsinstitutet (2018): att elevens styrkor och utvecklingsbehov ska synliggöras eller för att ge elever möjligheter att reflektera kring sitt lärande. De digitala proven underlättar denna återkoppling till elever i riskgruppen då visuella analyser av provutfallet samt svarsanalys kan tas fram.

Sambandet mellan resultat och betyg

Undersökningens syfte var också att undersöka huruvida det föreligger ett samband mellan elevernas resultat på DLS-proven och elevernas betyg. Den statistiska analysen, illustrerad i låddiagrammen visar att det finns en positiv samvarians mellan variablerna betyg och resultat. Ju högre poäng på DLS-proven elever har, desto högre betyg och tvärtom. Liknande positiva samband är funna av Carlsson, (2011); Levlin, (2014); Rehn & Sedvalsson, (2012); Helleklev (2013); Aldrin m.fl. (citerad i Johansson, 2016); Rudenius Sandin (citerad i Johansson, 2016) & Jacobsson (2014). Dessa studier inklusive denna pekar på att standardiserade tester för läs- och skrivutveckling har delvis ett predicerande och diagnostiskt värde när det gäller elever som inte bara kommer att ha svårigheter med sin läs- och skrivutveckling utan som kommer att ha svårt att bemästra skolans kunskapskrav i allmänhet. Att ett enstaka psykometriskt test kan ha en tydlig samvarians med ett mer komplext mått som betyg som innehåller många formativa och summativa bedömningar gjorda under en längre tid är intressant.

Låddiagrammen visar att det i synnerhet är elever med Stanine-värde under 4 som ligger i riskzonen för läs- och skrivsvårigheter och lägre betyg från F till D. Elever med resultat över 4 tenderar att få högre betyg från C-A.

Det positiva sambandet mellan elevernas prestationer på DLS-proven och elevernas betyg är stabilt oberoende av vilket prov vi använder och vilket ämne vi jämför med, vilket antyder att möjligen bara ett test skulle behöva användas i diagnostiskt syfte.

DLS-proven har en särskild ställning i skolan i det avseende att skolor väljer att köpa in och använda dem och att resultaten stannar på den enskilda skolan, till skillnad från nationella prov (SOU 2016:59) och i motsats till internationella kunskapsmätningar (Lundahl m.fl., 2016). De bör inte heller fungera betygsgrundande eller summativt (Järpsten & Taube, 2018). Därför får en formativ användning av DLS inte en funktion av ”upprepad testning” eller ”teaching to the test” som Klenowski, (2009), Chappuis, (2005) samt Hume & Coll (2009) kritiserar vid formativ användning av summativa prov. DLS-proven kan inte heller jämföras med s. k *benchmark assessment*, kommersiella standardiserade prov som skolor i USA använder under terminerna för att kartlägga vilka elever som inte kommer att klara de standardiserade slutproven. Summativa prov har kritiserats för att ignorera den sociala och historiska kontext i vilken bedömningen görs enligt sociokulturell teoribildning (James, 2012; Jönsson, 2018). Resultatet från en läs- och skrivscreening utan krav på publicering av resultat eller redovisning till extern aktör kan användas formativt med avseende på den kontext som skolan eller klasserna utgör. Att skolans personal involveras i genomförandet av proven och i vissa fall av rättningen ökar kopplingen till den sociala, miljömässiga och kulturella kontexten skolans personal och elever verkar i, vilket underlättar analys och formandet av undervisning.

Att vid genomförande av DLS-proven inspireras av det screeningförfarande som förekommer inom *Response to Intervention* (Gersten m.fl., 2008) skulle kunna medföra en relevant formativ användning av proven på skolnivå. Inom *RTI* används screeningresultat både för att fånga upp elever i behov av stöd och för att utvärdera hur effektiv den ordinarie undervisningen varit. *RTI* har potential att förena summativ och formativ bedömning, samt koppla den till systematiskt kvalitetsarbete på skol- och systemnivå, men samtidigt behöver tillvägagångssättet anpassas till svenska förhållanden.

Hinder och möjligheter för DLS-provens användande

Att förstå syftet med ett prov och veta hur resultaten ska användas är nödvändigt för alla i skolan som arbetar med bedömning (Wikström, 2013). Det konstateras att kartläggningar är vanligt, men att uppföljning och pedagogisk konsekvens av resultaten saknas (Myrberg & Lange, 2006). Det har visat sig i flera intervjustudier att DLS-proven ses som något som enbart berör specialpedagogen för hjälp till elever som kommit efter i sin läs- och skrivutveckling. Få lärare berörs eller använder resultaten som upplevs underutnyttjade i flera studier (Wiktorrell, 2016; Bäckman, 2011; Björnsson & Hellström, 2017; Bylund, 2011).

En tanke som förs fram av Jönsson (2018) är att gapet mellan psykometriska, tekniskt komplicerade kunskapsmätningar och bedömning i klassrummet har ökat. Baird m fl. (2017) menar också att den formativa bedömningen på klassrumsnivå står långt ifrån provkonstruktörers exakta och detaljerade vetenskap. DLS-proven kan av lärare utan erfarenhet av psykometriska test uppfattas som en ”svart låda”, svår att förstå sig på och använda.

Därtill förknippas psykometrin med beteendevetenskap på så sätt att den kunskap man mäter i DLS är uppdelad i små enheter och testas utanför sin meningsfulla kontext (James, 2008; Crossouard & Pryor, 2012). Beteendevetenskapen har fått ge vika för en sociokulturell syn på bedömning de senaste decennierna, med idéer om att kunskap konstrueras i mötet mellan människor och påverkas av individen, miljön och en mängd andra faktorer (Hirsh & Lindberg, 2015; James, 2012). Det sociokulturella synsättet har också givit avtryck i vår nuvarande läroplan med öppna, samhällsrelevanta betygskriterier som lärare ska kunna tolka och anpassa till varje elev, tvärtemot det normrelaterade betygssystemet (Jönsson, 2018; Wedman, 2003). DLS-proven kan verka främmande och oflexibla för dagens lärare som till viss del redan genomför formativ bedömning på klassrumsnivå, eftersom läroplanen i stor utsträckning kräver bedömning av kvalitén i elevernas prestationer och inte bara på deras kunskaper (Jönsson, 2018). Innehållet i DLS-proven kan förfalla kunskapsmässigt snävt i jämförelse med de breda beskrivningar av komplexa kunskaper och färdigheter i läroplanen. Proven ses mer kopplade till den kognitiva synen på att kunskap i vilken lärare överför kunskap till eleven, i en särskild ordning från det enkla och till det komplexa (Bloom, 1956).

Lärare kan också erfara att resultatet som presenteras i form av ett Stanine-värde som eleven kan jämföra med klasskamraternas inte är återkoppling som främjar lärande. Istället bör

relevant återkoppling relatera till en särskild uppgift eller till ett prov och beskriva hur eleven kan utveckla de färdigheter och kunskaper som bedömningen visat saknas (Black & Wiliam, 1998; Black m.fl., 2003; Hattie, 2009; Korp, 2011; Lundahl, 2014; Jönsson, 2018). Det är dock möjligt att med en kvalitativ analys av elevens provsvar göra en sådan formativ återkoppling till eleverna. Resultaten för skolan visar dock att detta inte görs ens för de allra svagaste eleverna, vilket är ett problem.

En slutsats som kan dras är att ramarna och förutsättningar för en formativ användning av DLS-resultaten behöver stärkas på skolan annars finns risk för att formativ bedömning används på ytligt sätt (Hirsh & Lindberg, 2015) och att tester görs utan pedagogiska konsekvenser. Hagtvét m. fl. (2016) betonar vikten av att kompetent skolpersonal med en fördjupad kunskap om läs- och skrivscreeningar kan leda arbetet med att tolka resultaten och använda dem i undervisningen. Specialpedagoger och speciallärare är de bäst lämpade för att driva detta arbete i den egna skolkontexten.

6.2 Metoddiskussion

Stukát (2011) anser att nackdelen med kvantitativa studier (fallstudier) är att de kan bli för generella, men saknar djup. För att få en ökad förståelse för lärarens eller skolledarens upplevelse rörande DLS-proven och dess resultat i relation till formativ bedömning hade intervjuer kunnat genomföras, men av tidsskäl gjordes detta inte. Värdefullt hade också varit att ta reda på mer om hur elever upplever DLS-provet, genom intervjuer eller eventuellt en enkät i anslutning till provtagningen. Hirsh & Lindberg (2015) antyder att elevperspektivet saknas när de har undersökt svenska och internationell forskning om formativ bedömning.

Att söka samband eller korrelation är en vanlig metod inom forskning (Stukát, 2011). Inom utbildning existerar mängder av samband och det är viktigt att skilja mellan samband som har praktisk betydelse för undervisningen, och samband som inte har det (Elbro & Poulsen, 2016).

Ett sätt att avgöra sambandets styrka är med Pearsons korrelationskoefficient (r) som också brukar kallas korrelationen och som är ett mått på hur perfekt sambandet är (Elbro & Poulsen, 2016). I denna undersökning låg korrelationen, alltså sambandet mellan alla tester och alla ämnen ungefär på 0,5, vilket anses vara ett medelstarkt till starkt samband. Det perfekta sambandet har korrelationen 1 och inget samband alls har korrelationen 0. Beräkningen av korrelationens styrka riskerar dock att inte ge korrekta resultat, eftersom de båda mätningarna som jämförs med varandra i undersökningen (DLS Resultat och betyg) skiljer sig åt på olika sätt. Styrkan på korrelationen blir tydligast om det är mätningar med liknande antal uppgifter som jämförs, vilket inte är fallet. Dessutom påverkas mätningen av korrelationens styrka negativt av att vissa elever får så kallade extrema resultat till exempel en elev med F i betyg kan få ett högt resultat på ett eller flera av testerna (Elbro & Poulsen, 2016). Korrelationer bör därför hellre visualiseras med hjälp av till exempel ett låddiagram som kan illustrera hur resultaten fördelar sig på alla elever.

Viktigt att komma ihåg är att korrelationen som låddiagrammen har visat bara är ett talsamband eller så kallat statistiskt samband där en variabel (DLS-resultat) och en annan variabel (betyg) samvarierar så att låga värden i den ena förekommer tillsammans med låga

värden i den andra variabeln. Det är dock inte säkert att det finns något orsakssamband eller kausalitet mellan Stanine-värde och betyg, utan det kan vara andra bakomliggande faktorer som påverkar orsakssambandet. Klapp (2016) har kommit fram till att betygen inte enbart mäter kunskap och färdigheter, utan att betygssättningen också påverkas av uppvisad motivation, könstillhörighet och socioekonomisk bakgrund. Ett betyg grundar sig dessutom på många olika bedömningstillfällen, både formativa och summativa, medan DLS-provet enbart görs vid ett tillfälle. Vid ett enstaka provtillfälle kan resultatet påverkas av att eleverna är nervösa, har sovit dåligt eller av olika omständigheter har svårt att visa sina kunskaper (Klapp, 2016). Trots svårigheterna med att avgöra reliabilitet och validitet med avseende på elevernas betyg, får vi ändå anta att betygssättning på skolan skett enligt kraven i Lgr11 (Skolverket, 2018) och att lärarna haft tillfälle att inhämta bedömningar på ett sätt som gör betygen tillförlitliga och giltiga. Det finns dock en risk att elever kan underpresterar på DLS-proven eftersom proven inte är summativa och inte ligger till grund för betyg.

6.3 Studiens kunskapsbidrag

En av studiens kontributioner är att alla fyra läs- och skrivprov har bekräftats som ett diagnostiskt instrument för att upptäcka inte bara elever som riskerar problem med läs- och skrivförmågan, utan också de som riskerar att inte nå grundskolans mål i ämnena svenska, matematik och historia/samhällskunskap, på den undersökta skolan. Detta indikerar att något av de fyra proven enskilt skulle kunna användas diagnostiskt, för att undvika alltför omfattande testning. Standardiserade läs- och skrivprov är således lämpliga att användas i en formativ bedömningsprocess med fokus på att forma lärandet och inte eleverna (Hirsh, 2017), men också för att identifiera elever vars läsning och skrivning är i behov av utveckling. När resurser är begränsade behöver de instrument som används för att ge information om elevers nivå i högre grad användas för att utveckla elevers lärande. Det framkommer att till exempel att resultatet för elever med svenska som andra språk eller hur klasser presterar kan ligga till grund för undervisningsjusterande åtgärder på grupp- och skolnivå. Elevhälsans förebyggande arbete underlättas av att DLS-resultaten publiceras tidigt på terminen. Med avseende på det positiva samband som existerar mellan DLS-provens resultat och elevernas betyg i skilda ämnen bör elevhälsan prioritera en uppföljning av elever med Stanine-värde 1–3 där en plan omkring elevens utveckling ska ingå.

6.4 Förslag till vidare forskning

Vid en litteraturgenomgång kan det konstateras att DLS-proven är ett relativt utforskat område, trots sin vida användning. En fallstudie har en begränsad generaliserbarhet (Yin, 2007) och det torde vara angeläget att fortsätta studera detta område både med kvantitativa och kvalitativa utgångspunkter på fler skolor. Särskilt i de fall där en hel kommun genomför proven kan det finnas goda möjligheter att utforska både resultatet över tid, hur resultaten används på individ-,grupp- och skolnivå samt inställningen till proven hos skolpersonal och elever. I denna undersökning har jag bland annat granskat DLS-resultaten i relation till ämnena svenska och svenska som andra språk. Många fler elever än de som läser svenska som andra språk har ett annat modersmål än svenska på den aktuella skolan. Att jämföra alla dessa elevers resultat med de som enbart har svenska som hemspråk vore intressant för att i större utsträckning få information om vad dessa elever befinner sig i sin läsning och skrivning, som sedan kan ligga till grund för en förändring av undervisningen. Studier som fokuserar på uppföljning av elever med låga screeningresultat är också angeläget att genomföra eftersom användning av läs- och skrivscreeningar rekommenderas också upp i åldrarna av läsforskare och intresseorganisationer (Hagtvét m. fl, 2016; Roe, 2015; Fridolfsson, 2015; Lundberg, 2010; Dyslexiföreningen, 2017). Ytterligare kunskap om hur *Response to Intervention* förenar screeningresultat med ett proaktivt arbete på individ-, grupp- och skolnivå för att minimera antalet elever som inte når målen är önskvärt och skulle kunna inspirera till aktionsforskning och jämförande studier mellan länder.

7 Referenslista

- Aldrin, K., Andersson, S-E & Holm, L. (1997). *Läst, men ej förstått – läs- och stavningsförmågans utveckling hos elever på en högstadieskola i Mellansverige*. (Uppsats). Linköping: Linköpings Universitet: Läspedagogiska institutet EMIR.
- Baird, J., Hopfenbeck, T., Newton, P., Stobart, G., & Steen-Utheim, A. (2014). *State of the field review: Assessment and learning*. Oslo: Norwegian Knowledge Centre for Education.
- Baird, J., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and Learning: Fields Apart? *Assessment in Education: Principles, Policy & Practice* 24.3: 317-350.
- Bennett, R., E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18:1, 5-25, doi: [10.1080/0969594X.2010.513678](https://doi.org/10.1080/0969594X.2010.513678)
- Björnsson, M., & Hellström, L. (2017). *Möjligheter och utmaningar med läsförståelsetest - speciallärares upplevelser av screening* (Examensarbete). Stockholm: Stockholms Universitet: Specialpedagogiska Institutionen. Hämtad 19-05-15 <http://su.diva-portal.org/smash/get/diva2:1153063/FULLTEXT01.pdf>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 1-74. doi: <https://doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. (2009). Developing the Theory of Formative Assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21. (1), 5-31.
- Black, P., & Wiliam, D. (2018). Classroom Assessment and Pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25, (6), 551-575.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it into practice*. Maidenhead: Open University Press.
- Blomqvist, P. (2018). *Samtal om skrivbedömning: lärares normer, beslut och samstämmighet*. (Doktorsavhandling, Doktorsavhandlingar i språkdidaktik :1) Stockholm: Stockholms universitet. Hämtad från <http://su.diva-portal.org/smash/record.jsf?pid=diva2%3A1197840&dswid=607>
- Bloom, B., S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation*. New York: McGraw-Hill Inc.
- Boström, E. (2017). *Formativ bedömning: en enkel match eller en svår utmaning? Effekter av en kompetensutvecklingssatsning på lärarnas praktik och på elevernas prestationer i*

- matematik* (Doctoral thesis, Umeå Studies in Science and Technology Education, 96)
 Umeå: Umeå universitet. Hämtad från <http://umu.diva-portal.org/smash/record.jsf?pid=diva2%3A1096034&dswid=3209>
- Broadfoot, P., Daugherty, R., Gardner, J., Gipps, C., Harlen, W., James, M., & Stobart, G. (1999). *Assessment for learning: beyond the black box*. Nuffield Foundation and University of Cambridge. DOI: 10.13140/2.1.2840.1444
- Broadfoot, P., & Black, P. (2004). The first ten years of Assessment in Education. *Assessment in Education: Principles, Policy & Practice*, 11(1), 7-26.
- Bryman, A. (2011). *Samhällsvetenskapliga metoder*. Malmö: Liber.
- Bylund, M. (2018). *Screening av elevers språk-, läs- och skrivutveckling. En hermeneutisk studie av specialpedagogers erfarenheter av screeningar* (Examensarbete). Karlstad: Fakulteten för humaniora och samhällsvetenskap. Hämtad från <http://kau.diva-portal.org/smash/record.jsf?pid=diva2%3A1222858&dswid=-8793>
- Bäckman, K. (2011). *Gymnasieelevers läsförståelse. En grupp pedagogers åsikter kring screeningtest och läsförståelseundervisning* (Magisteruppsats). Göteborg: Göteborgs Universitet: Institutionen för pedagogik och specialpedagogik. Hämtad från <https://gupea.ub.gu.se/handle/2077/27005?locale=sv>
- Carlsson, M. (2017). *Läsförståelse och stadiövergång - en jämförande textanalys av Nationellt prov i svenska för år 6 och DLS screening för år 7 och 8*. (Master's thesis). Malmö: Malmö Universitet: Institutionen för lärande och samhälle. Hämtad från <http://muep.mau.se/handle/2043/23507>
- CERI. (2005). *Formative Assessment: Improving learning in secondary classrooms*. Paris: OECD Publishing.
- Chappuis, S. (2005). Is Formative Assessment Losing Its Meaning? *Education Week*, 24 (44), p 38.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis*. New Jersey: Pearson Education.
- Crossouard, B., & Pryor, J. (2012). How Theory Matters: Formative Assessment Theory and Practices and Their Different Relations to Education. *Stud Philos Educ*, 31, 251–263. doi: DOI 10.1007/s11217-012-9296-5
- Dixon, D. D., & Worrell, F. C. (2016). Formative and Summative Assessment in the Classroom. *Theory Into Practice*, 55 (2), 153-159. doi:10.1080/00405841.2016.1148989

- Djuvfeldt, Gunilla & Wedman, Ingemar. (2007). *Betyg i praktiken: lärares och arbetsgivares syn på betyg*. Gävle: Högskolan i Gävle.
- Dyslexiföreningen. (2017). *Modell för utredning av läs- och skrivsvårigheter*. Hämtat 2019-02-15 från https://www.dyslexi.org/sites/default/files/utredningsmodell_2017.pdf
- Ejvegård, R. (2009). *Vetenskaplig metod*. Lund: Studentlitteratur AB.
- Elbro, C., & Poulsen, M. (2016). *Utvärdera din undervisning: värdera och förstå statistik och evidens*. Stockholm: Natur & kultur.
- Fridolfsson, I. (2015). *Grunderna i läs- och skrivinlärning*. Lund: Studentlitteratur.
- Gersten, R., Compton, D., Connor, C.M., Dimino, J., Santoro, L., Linan-Thompson, S., and Tilly, W.D. (2008). *Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide*. (NCEE 2009-4045). Hämtad från What Works Clearinghouse: <https://ies.ed.gov/ncee/wwc/PracticeGuide/3>
- Gipps, C. V. (1994). *Beyond testing. Towards a theory of educational assessment*. London: The Falmer Press. Hämtad från <https://epdf.tips/beyond-testing-towards-a-theory-of-educational-assessment.html>
- Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24, 355- 392. Hämtad från <https://doi-org.ezproxy.ub.gu.se/10.3102/0091732X024001355>
- Gough, P. B., Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, 7, 6–10. doi:10.1177/074193258600700104
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, 18(8),519-521. doi: <http://dx.doi.org.ezproxy.ub.gu.se/10.1037/h0049294>
- Gustafsson, J.-E., Cliffordson, C., & Erickson, G. (2014). *Likvärdig kunskapsbedömning - i och av den svenska skolan*. SNS Utbildningskommission. Stockholm: SNS Förlag.
- Gustavsson, A., Måhl, P., & Sundblad, B. (2012). *Betygssättning – en handbok*. Stockholm: Liber AB.
- Hagtvet, B.E., Frost, J., & Refsahl, V. (2016). *Intensiv läsinlärning: dialog och bemästrande när läsningen har låst sig*. Stockholm: Studentlitteratur.
- Hattie. J. (2009). *Visible Learning*. London: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112. Hämtad <https://doi-org.ezproxy.ub.gu.se/10.3102/003465430298487>

- Helleklev, C. (2013). *Läsförståelse i gymnasiet - en självklarhet?* (Magisteruppsats). Halmstad: Högskolan i Halmstad: Institutionen för utbildningsvetenskap. Hämtad från <http://hh.diva-portal.org/smash/record.jsf?pid=diva2%3A633161&dsid=90>
- Hirsh, Å., & Lindberg, V. (2015). *Formativ bedömning på 2000-talet - en översikt av svensk och internationell forskning*. Stockholm: Vetenskapsrådet. Hämtad från https://www.vr.se/download/18.2412c5311624176023d25aca/1529480531147/Formativ-bedomning-svensk-o-internationell-forskning_VR_2015.pdf den 28 02 2019
- Hirsh, Å. (2017). *Bedömning i skolan – vad och varför?* Stockholm: Skolverket. Hämtad från <http://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-37821>
- Hogrefe Förlag AB. (2019a). *DLS för skolår 7-9 samt 1 på gymnasiet*. Hämtad 2019-05-06 från <https://www.hogrefe.se/Skola/Las--och-skrivdiagnostik1/Screening-NY/DLS-7-9-och-ar-1-i-gymnasiet---reviderade-delprov-samt-handledning/>
- Hogrefe Testsystem. (2019b). Hämtad 2019-05-09 från <https://www.hogrefe-online.com/HTSEnvironment/main>
- Hoover, W. A., & Tunmer, W. E. (2018). The Simple View of Reading: Three Assessments of Its Adequacy. *Remedial and Special Education*, 39(5), 304–312. Doi:[10.1177/0741932518773154](https://doi.org/10.1177/0741932518773154)
- Hume, A., & Coll, R. K. (2009). Assessment of learning, for learning, and as learning: New Zealand case studies. *Assessment in Education: Principles, Policy & Practice*, 16(3), 269-290. Doi: [10.1080/09695940903319661](https://doi.org/10.1080/09695940903319661)
- Jacobson, C. (2014). *Läskedjor - 2*. Stockholm: Hogrefe Psykologiförlaget AB.
- James, M. & Lewis, J. (2012). Assessment in Harmony with our Understanding of Learning Assessment in Harmony with our Understanding of Learning: Problems and Possibilities. I J. Gardner (Red.), *Assessment and Learning, 2nd edition*. London: Sage.
- James, M. (2006). Assessment, teaching and theories of learning. I J. Gardner (Red.), *Assessment and Learning* (s. 47-60). London: Sage.
- James, M. (2008). Assessment and learning. I S. Swaffield (Red.), *Unlocking assessment. Understanding for reflection and application* (s. 20–35). Abingdon: Routledge.
- Johansson, M.-G. (2016). *Klassdiagnoser i läsning och skrivning för högstadiet och gymnasiet - LS*. Stockholm: Hogrefe Psykologiförlaget AB.
- Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*, 78(4), 243-255. doi: [10.1037/0022-0663.78.4.243](https://doi.org/10.1037/0022-0663.78.4.243)

- Järpsten, B. (2002). *DLSHandledning för skolår 7-9 och år 1 i gymnasiet*. Stockholm: Hogrefe Psykologiförlaget AB.
- Järpsten, B., & Taube, K. (2018). *DLS För skolår 7-9 och år 1 i gymnasiet. Rättstavning, Ordförståelse, Lëshastighet och Läsförståelse*. Stockholm: Hogrefe Psykologiförlaget.
- Jönsson, A., & Balan, A. (2014). *Bedömning för lärande - en vägledning utifrån aktuell forskning*. (Forskning i korthet, 2014:2). Lund: FoU Skola/kommunförbundet Skåne.
- Jönsson, A. (2017). *Lärande Bedömning*. Malmö: Gleerups Utbildning.
- Jönsson, A. (2018). *Prov eller bedömning?*. Malmö: Gleerups Utbildning.
- Klapp, A. (2016). Vad mäter betygen?. I Lundahl, C. & Folke-Fichtelius, M. (red.) *Bedömning i och av skolan: praktik, principer, politik*. Lund: Studentlitteratur.
- Klenowski, V. (2009). Assessment for learning revisited: an Asia-Pacific perspective. *Assessment in Education: Principles, Policy & Practice*, 16(3), 263-268. Doi: 10.1080/09695940903319646.
- Korp, H. (2011). *Kunskapsbedömning. Vad, hur och varför?* Stockholm: Skolverket. Hämtad från <https://www.skolverket.se/getFile?file=2666>
- Levlin, M. (2014). *Lässvårigheter, språklig förmåga och skolresultat i tidiga skolår: en undersökning av 44 elever i årskurs 2 till 3*. Diss. Umeå : Umeå universitet, 2014. Umeå. Hämtad från <http://www.diva-portal.org/smash/get/diva2:763839/FULLTEXT01.pdf>
- Lundberg, I. (2010). *Läsningens psykologi och pedagogik*. Stockholm: Natur & kultur.
- Lundahl, C, Hultén, & Tveit, S. (2017). Betygssystem i internationell belysning. Skolverket.
- Lundahl, C. (2014). *Bedömning för lärande*. Lund: Studentlitteratur AB.
- Myrberg, M. (2003). *Konsensus om skolans insatser för att förebygga och möta läs- och skrivsvårigheter*. Stockholm: Specialpedagogiska Institutet och Lärarhögskolan i Stockholm
- Myrberg, M. & Lange, A-L (2006). *Identifiering, diagnostik, samt specialpedagogiska insatser för elever med Läs- och skrivsvårigheter*. Rapport från Konsensusprojektet. Stockholm: Specialpedagogiska Institutet och Lärarhögskolan i Stockholm.
- Nationellt centrum för svenska som andraspråk (2019). *Svenska som andra språk – Grundskola*. Hämtad 2019-05-14 från <https://www.andrasprak.su.se/grundskola/styrdokument/svenska-som-andraspr%C3%A5k-grundskola-1.82989>

- Patel, R., & Davidson, B. (2015). *Forskningsmetodikens grunder. Att planera, genomföra och rapportera en undersökning*. Lund: Studentlitteratur.
- Pettersson, A (2010): *Bedömning för lärande – en grund för ökat kunskande*. Rapport 331:3 Januari 2010. Stockholm: Stiftelsen SAF i samarbete med Lärarförbundet.
- Nationalencyklopedin (2019). *Screening*. Hämtad 2019-03-15 från <https://www-ne-se.ezproxy.ub.gu.se/uppslagsverk/encyklopedi/l%C3%A5ng/screening>
- Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford review of Education*, 34(1), 1–20. DOI: [10.1080/03054980701476386](https://doi.org/10.1080/03054980701476386)
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science*, 28(1), 4-13.
- Rehn, C., & Sedvallson, M. (2012). *Diagnostiskt läs- och skrivprov - ett verktyg för att identifiera elever som riskerar att ej nå målen i svenska, år 3?* (Master's thesis). Falun: Högsolan i Dalarna: Institutionen för didaktik. Hämtad från <https://www.uppsatser.se/om/diagnostiskt+l%C3%A4s+och+skrivprov/>
- Reichenberg, M. (2008). *Vägar till läsförståelse*. Stockholm: Natur & Kultur.
- Resnick, B., L., & Schantz, F. (2017). Testing, teaching, learning: who is in charge?, *Assessment in Education: Principles, Policy & Practice*, 24(3), 424-432, DOI: [10.1080/0969594X.2017.1336988](https://doi.org/10.1080/0969594X.2017.1336988)
- Roe, A. (2015). *Läsdidaktik*. Malmö: Ekelund/Gleerups utbildning.
- Roos, B., & Hamilton, D. (2005). Formative assessment: A cybernetic viewpoint. *Assessment in Education: Principles, Policy & Practice*, Vol. 12, (1) 7-20.
- Rudénius Sandin, B. (1997). *Språklig förmåga hos elever i årskurs 8. En undersökning i en svensk kommun*. 60-poängsuppsats. Institutionen för pedagogik och specialmetodik. Lärarhögskolan: Lunds Universitet, Malmö.
- Sadler, R. D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Selghed, B. (2004). *Ännu icke godkänt. Lärare sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Diss., Malmö: Malmö Högskola.
- SBU. (2014). *Dyslexi hos barn och unga: Tester och insatser. En systematisk litteraturöversikt*. (SBU-rapport nr 225). Stockholm: Statens beredning för medicinsk utvärdering.
- Scriven, M. (1967). The methodology of evaluation. i R. W. Tyler, R. M. Gagné, & M. Scriven, *Perspectives of curriculum evaluation* (s. 39-83). Chicago: Rand McNally.

- Shepard, L. (2001). *The role of classroom assessment in teaching and learning*. I In. V. Richardson (red.) *Handbook of research on teaching*, s 1066-1101. Washington DC: AERA.
- SFS 2010:800. *Skollag*. Stockholm: Utbildningsdepartementet.
- Skolforskningsinstitutet (2018). *Feedback i skrivundervisningen. Systematisk översikt*. (2018:01). Solna: Skolforskningsinstitutet.
- Skolverket. (2011a). *Kunskapsbedömning i skolan. Praxis, begrepp, problem och möjligheter*. Stockholm: Skolverket. Hämtad från https://www.skolverket.se/sitevision/proxy/publikationer/svid12_5dfee44715d35a5cdfa2899/55935574/wtpub/ws/skolbok/wpubext/trycksak/Blob/pdf2660.pdf?k=2660
- Skolverket. (2011b). *Kunskapsbedömning – vad, hur och varför*. Stockholm: Skolverket. Hämtad från: <https://www.skolverket.se/publikationer?id=2666>
- Skolverket. (2016). *PISA 2015 – 15-åringars kunskaper i naturvetenskap, läsförståelse och matematik*. Stockholm: Skolverket.
- Skolverket. (2018a). *Skolverkets allmänna råd. Betyg och betygssättning*. Stockholm Skolverket. Hämtad från <https://www.skolverket.se/publikationsserier/allmanna-rad/2018/allmanna-rad-om-betyg-och-betygssattning>
- Skolverket. (2018b). *Forskningsbaserat arbetssätt för ökad kvalitet i skolan*. Hämtad från <https://www.skolverket.se/skolutveckling/forskning-och-utvarderingar/forskningsbaserat-arbetssatt/forskningsbaserat-arbetssatt-for-okad-kvalitet-i-skolan>
- Skolverket. (2018c). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011: reviderad 2018. (5:e uppl.)* Hämtad från <https://www.skolverket.se/undervisning/grundskolan/laroplan-och-kursplaner-for-grundskolan/laroplan-lgr11-for-grundskolan-samt-for-forskoleklassen-och-fritidshemmet>
- Skolverket. (2019a). *Formativ bedömning*. Hämtad från <https://www.skolverket.se/undervisning/kallsidor/formativ-bedomning>
- Skolverket. (2019b). *Formativ bedömning stärker lärandet*. Hämtad 2019-03-08 från <https://www.skolverket.se/skolutveckling/forskning-och-utvarderingar/forskning/formativ-bedomning-starker-larandet>
- Skolverket. (2019c). *Bedömningsportalen*. Hämtad från Skolverket: Bedömningsportalen: <https://bp.skolverket.se/>

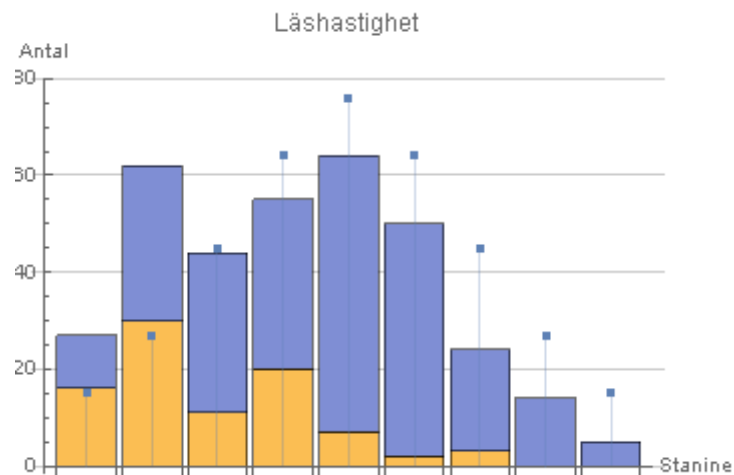
- Skolverket. (2019d). Nationella prov. Hämtad från Skolverket: <https://www.skolverket.se/for-dig-som-ar.../elev-eller-foralder/betyg-och-nationella-prov/nationella-prov>
- Skolverket. (2019e). *PISA – en studie i elevers kunskaper i matematik, naturvetenskap och läsförståelse*. Hämtad från Skolverket: <https://www.skolverket.se/skolutveckling/forskning-och-utvarderingar/internationella-jamforande-studier-pa-utbildningsområdet>
- Skolverket. (2019f). *Kartläggning i förskoleklass*. Hämtad från Skolverket: <https://www.skolverket.se/undervisning/forskoleklassen/kartlaggning-i-forskoleklassen>
- Socialstyrelsen/Skolverket. (2016). *Vägledning för elevhälsan*. Hämtad från Socialstyrelsen: <https://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/20394/2016-11-4.pdf>
- Sundsvalls kommun. (2016). *Plan för screening av Läs- och skrivförmåga*. Hämtad från Sundsvalls kommun: https://pedagogisundsvall.se/wp-content/uploads/2016/03/plan_for_screening_av_las_och_skrivformaga.pdf
- Stukát, S. (2011). *Att skriva examensarbete inom utbildningsvetenskap*. Lund: Studentlitteratur AB.
- Säljö, R. (2000). *Lärande i praktiken: ett sociokulturellt perspektiv*. Stockholm: Prisma.
- Taube, K., Fredriksson, U., & Olofsson, Å. (2015). *Kunskapsöversikt om läs- och skrivundervisning för yngre elever : Delrapport från SKOLFORSK-projektet*. Hämtad från <http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-101726>
- Täby Barn och Grundskolan. (2016). *Screeningplan. Förhållningssätt och verktyg för att skapa förutsättningar för en god läs-, skriv- och matematikutveckling*. Hämtad från <https://www.taby.se/globalassets/3.-dokument-per-dokumenttyp/information/naringsliv-och-foretagande/screeningplan.pdf>
- SOU 2016:59. *På goda grunder - en åtgärdsgaranti för läsning, skrivning och matematik: betänkande*. Stockholm: Wolters Kluwer.
- SOU 2017:35. *Samling för skolan Nationell strategi för kunskap och likvärdighet*. Stockholm: Wolters Kluwer.
- Vetenskapsrådet. (2017). *God forskningssed*. Hämtad från Vetenskapsrådet: https://www.vr.se/download/18.2412c5311624176023d25b05/1529480532631/God-forskningssed_VR_2017.pdf
- Vetenskapsrådet. (2015a). *Kunskapsöversikt om läs- och skrivundervisning för yngre elever*. Hämtat från Vetenskapsrådet: <https://www.vr.se/analys-och-uppdrag/vi-analyserar->

- och-utvärderar/alla-publikationer/publikationer/2015-06-10-kunskapsöversikt-om-las--och-skrivundervisning-for-yngre-elever.-delrapport.html
- Vetenskapsrådet. (2019). *Forskningsöversikt 2019*. Utbildningsvetenskap. Stockholm: Vetenskapsrådet. Hämtad från Vetenskapsrådet: https://www.vr.se/download/18.ad27632166e0b1efab472f/1552382329529/Forskningsoversikt-utbildningsvetenskap_VR_2019.pdf
- Vlachos, J. (2018). *Trust-Based Evaluation in a Market-Oriented School System*. (IFN Working Paper No. 1217). Stockholm: Research Institute of Industrial Economics.
- Västerås Stad. (2016). *Screeningplan 2016. Förhållningssätt och verktyg för att skapa förutsättningar för en god läs- och skrivutveckling*. Hämtad från Västerås Stad: <https://docplayer.se/19375100-Screeningplan-2016-forhallningssatt-och-verktyg-for-att-skapa-forutsattningar-for-en-god-las-och-skrivutveckling.html>
- Wallin, E. (1968). *Mot en pedagogisk teknologi*. I Wallin (red.). *Undervisning – konst eller teknik?* (s. 9-13). Stockholm: Almqvist & Wiksell.
- Wallin, E. (2005). The Rise and Fall of Swedish Educational Technology 1960-1980. *Scandinavian Journal of Educational Research*, 49 (5), 437-460.
- Wedman, Ingemar. (2003). Från skallmätning till VG. Om skolans betyg. I S. Selander (red.). *Kobran, nallen och majjen. Tradition och förnyelse i svensk skola och skolforskning*. (s. 307-316). Forskning i fokus, nr 12. Stockholm: Myndigheten för skolutveckling.
- Wiktorell, Paula. (2016). *Återkommande kartläggning av elevers läs- och skrivförmåga. Hur en systematisk uppföljning används i en kommun* (Master's thesis). Göteborg. Institutionen för pedagogik och specialpedagogik, Göteborgs Universitet. Hämtad från <https://gupea.ub.gu.se/handle/2077/45033>
- William, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? I C. A. Dwyer (Red.), *The future of assessment: shaping teaching and learning*. (s. 53-82). Mahwah, NJ: Lawrence Erlbaum Associates.
- William, D. (2009). *Assessment for learning – why, what and how?* University of London: Institute of Education.
- Wikström, C. (2013). *Konsten att göra bra prov. Vad lärare behöver veta om kunskapsmätning*. Stockholm: Natur & Kultur.
- Wikström, C. (2017). Proven i skolan. I A. Hult, A. Olofsson, A. (Red.), *Utvärdering Och Bedömning I Skolan: För Vem Och Varför?* Stockholm: Natur & Kultur.
- Yin, R. K. (2007). *Fallstudier: design och genomförande*. Stockholm: Liber AB.

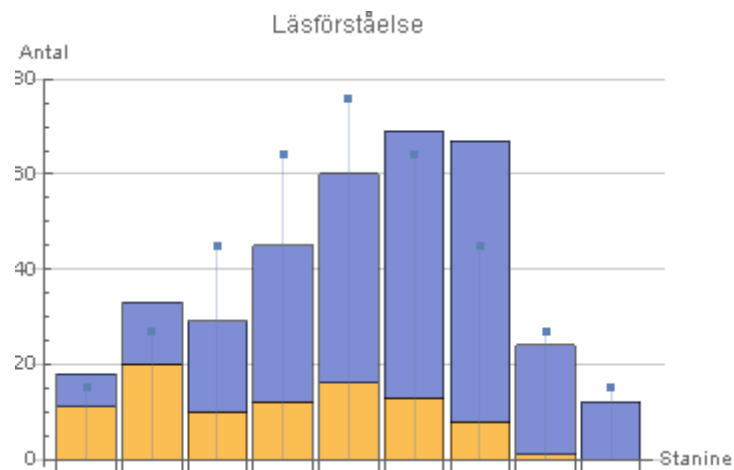
Bilaga 1

Skillnaden mellan elever som läser svenska och svenska som andra språk

Figurerna visar skillnaden i resultat från DLS-proven läshastighet och läsförståelse, för elever som läser svenska som andra språk och elever som läser svenska. På vågrät axel presenteras Stanine-värde från 1–9 och på lodrät axel de elever som fick respektive poäng.



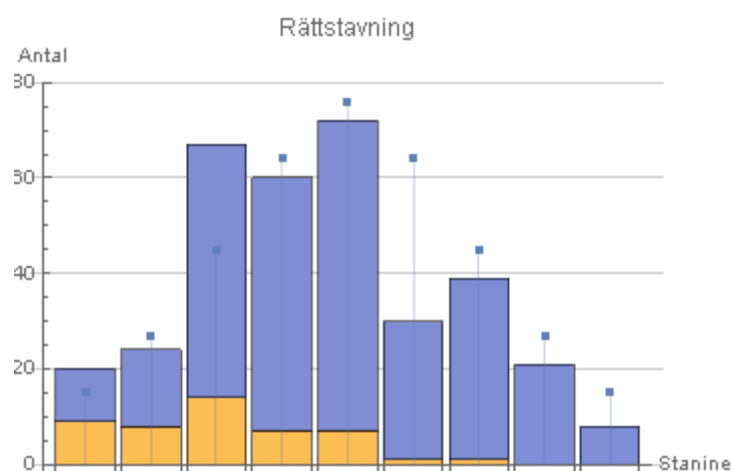
Figur 13. Skillnaden i resultat mellan elever som läser svenska och svenska som andra språk, vid provet läshastighet.



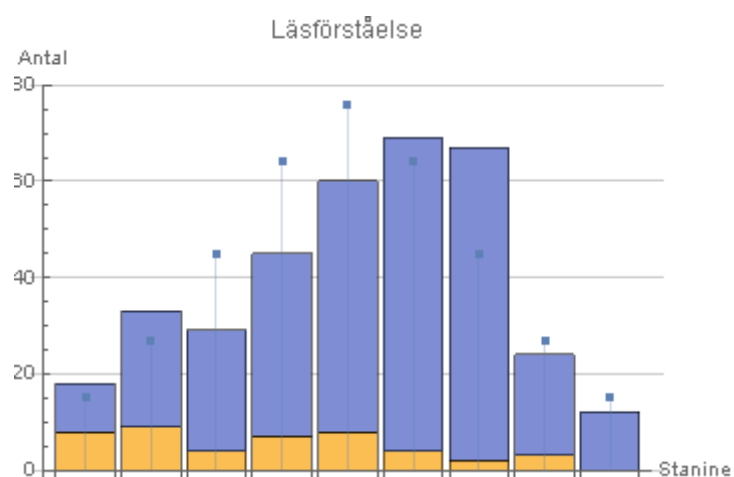
Figur 14. Skillnaden i resultat mellan elever som läser svenska och svenska som andra språk, vid provet läsförståelse.

Bilaga 2

Andel elever som får särskilt stöd



Figur 15. Andel elever som får särskilt stöd sett till resultatet i rättstavningsprovet

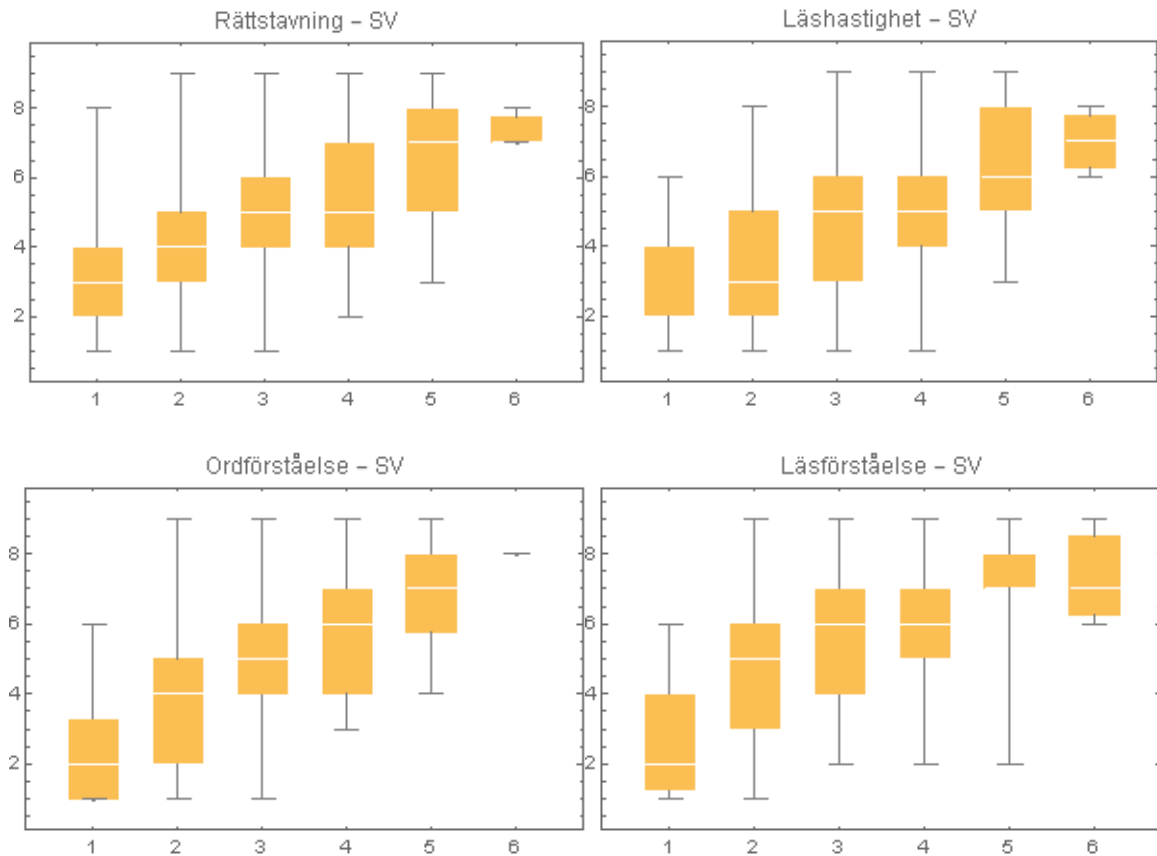


Figur 16. Andel elever som får särskilt stöd sett till resultatet i läsförståelseprovet

Bilaga 3

Sambandet mellan Stanine och betyg i svenska för alla fyra proven

Här nedan presenteras sambanden för alla fyra proven tillsammans för en bättre överblick. På vågrät axel återfinns betygen från F-A (1-6) och på lodrät axel Stanine-värden 1-9.

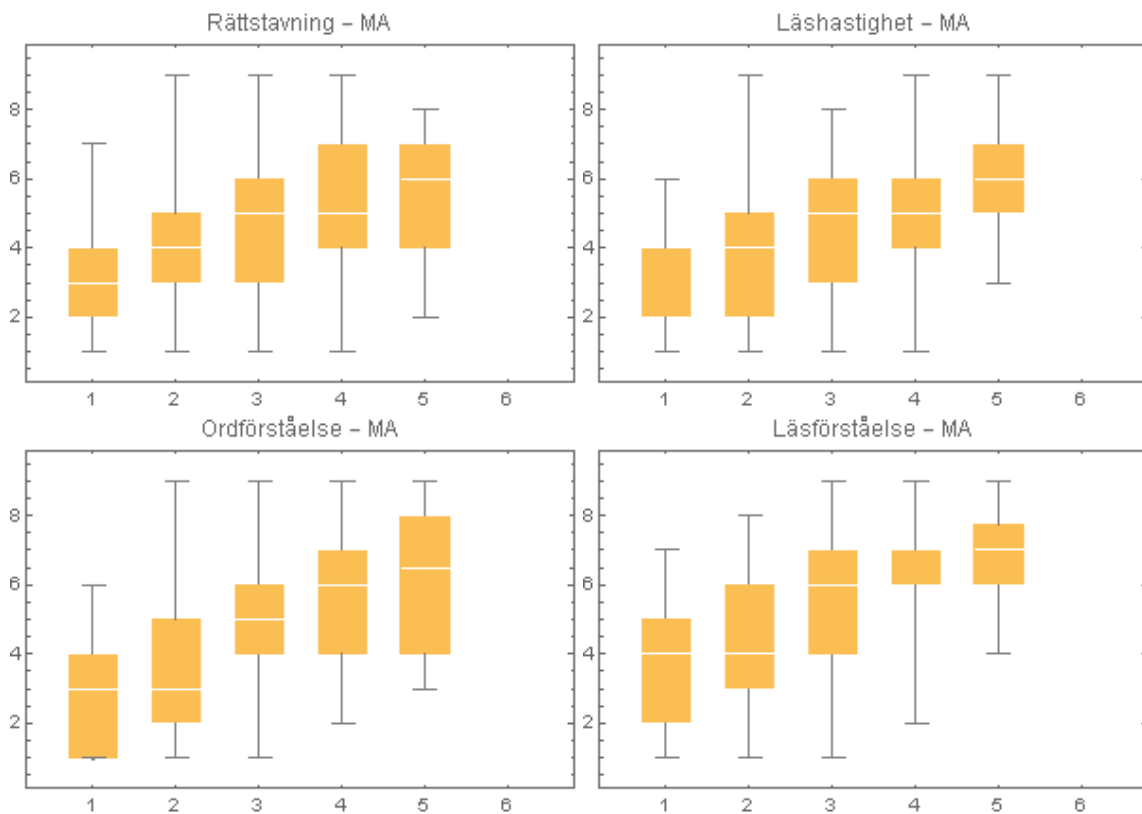


Figur 17. Sambanden för alla fyra proven och betyg i svenska och svenska som andra språk

Bilaga 4

Sambandet mellan Stanine och betyg i matematik för alla fyra proven

Här följer låddiagram som visar provens samband med betyget ämnet matematik. På vågrät axel återfinns betygen från F-A (1–6) och på lodrät axel Stanine-värden 1-9.

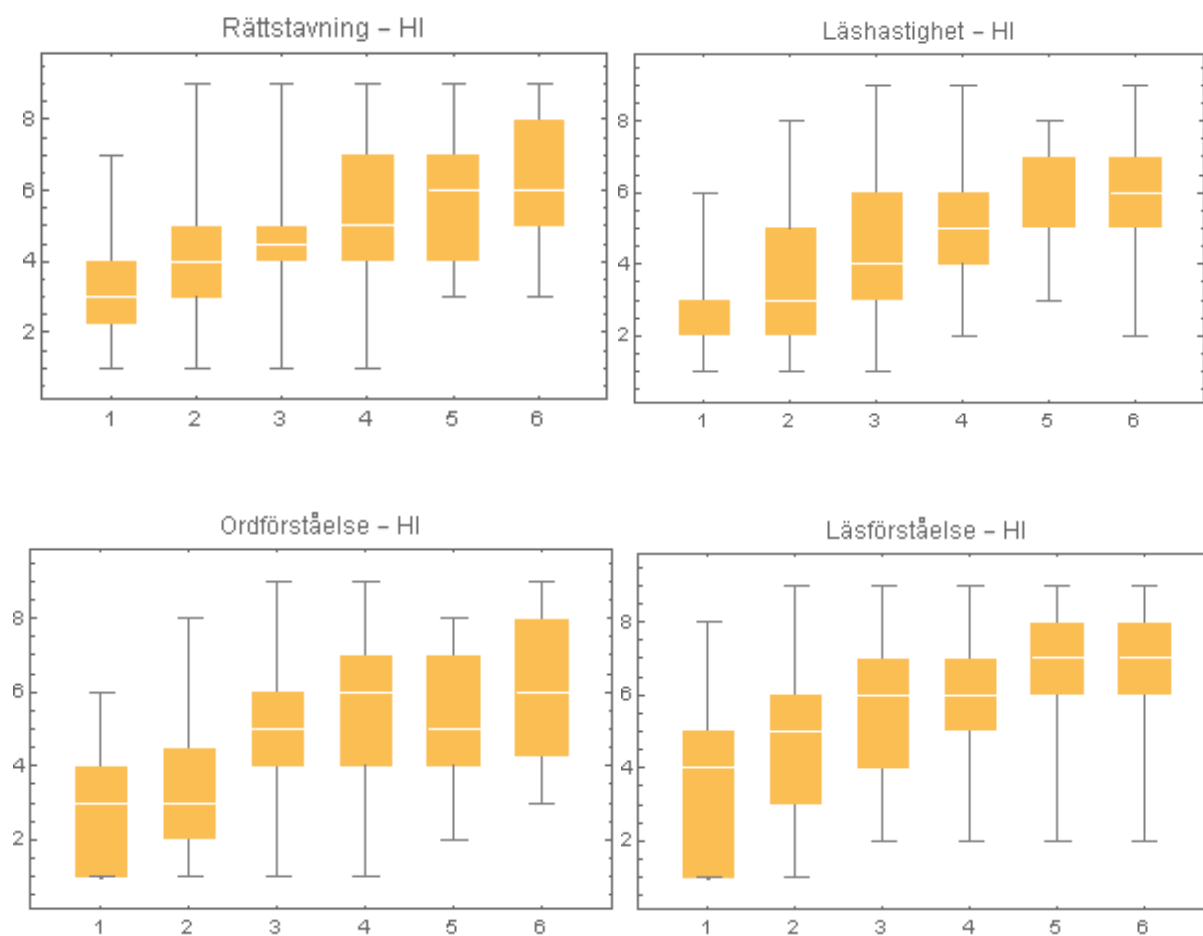


Figur 18. Sambanden för alla fyra proven och betyg i matematik

Bilaga 5

Sambandet mellan Stanine och betyg i historia och samhällskunskap för alla fyra proven

Till sist illustreras låddiagram över alla fyra proven korrelerat till betygen i historia och samhällskunskap.



Figur 19. Sambanden för alla fyra proven och betyg i matematik